Online Advertisement Allocation Under Customer Choices and Algorithmic Fairness

Xiaolong Li
National University of Singapore, oralxi@nus.edu.sg

Ying Rong
Shanghai Jiao Tong University, yrong@sjtu.edu.cn

Renyu Zhang
The Chinese University of Hong Kong, philipzhang@cuhk.edu.hk

Huan Zheng
Shanghai Jiao Tong University, zhenghuan@sjtu.edu.cn

August 27, 2023

Advertising is a crucial revenue source for e-commerce platforms and a vital online marketing tool for their sellers. In this paper, we explore dynamic ad allocation with limited slots upon each customer arrival for an e-commerce platform, where customers follow a choice model when clicking the ads. Motivated by the recent advocacy for the algorithmic fairness of online ad delivery, we adjust the value from advertising by a general fairness metric evaluated with the click-throughs of different ads and customer types. The original online ad-allocation problem is intractable, so we propose a novel stochastic program framework (called *two-stage target-debt*, TTD) that first decides the click-through targets then devises an ad-allocation policy to satisfy these targets in the second stage. We show the asymptotic equivalence between the original problem, the relaxed click-through target optimization, and the fluid-approximation (Fluid) convex program. We also design a debt-weighted offer-set (DWO) algorithm and demonstrate that, as long as the problem size scales to infinity, this algorithm is (asymptotically) optimal under the optimal first-stage click-through target. Compared to the Fluid heuristic and its re-solving variants, our approach has better scalability and can deplete the ad budgets more smoothly throughout the horizon, which is highly desirable for the online advertising business in practice. Finally, our proposed model and algorithm help substantially improve the fairness of ad allocation for an online e-commerce platform without significantly compromising efficiency.

Key words: Online Advertising Platform, Assortment Optimization, Algorithmic Fairness, Online Convex Optimization

1. Introduction

The past 10 years have witnessed the rapid growth of internet technology and smartphone penetration, both of which have driven online advertising to become an unprecedentedly enormous trillion-dollar industry[1] that has an enormous impact on the entire economy.

[1] Interactive Advertising Bureau report. See https://www.iab.com/insights/internet-advertising-revenue-fy2019-q12020/.

One important online advertising format is e-commerce advertising, which is designed to drive "top-of-tunnel" traffic to convert into product sales. For instance, Amazon Advertising provides "sponsored products,"[2] where advertisers pay Amazon to promote their products by listing the ads both within shopping results and on product pages (see Figure 1). The sponsored-product ads use the cost-per-click (CPC) mechanism, under which advertisers pay a fee to the platform when customers click their ads. Advertisers choose the campaign budgets and how much to bid per click. Amazon also allows advertisers to set the keywords and products so that the ad can be more efficiently matched with customer queries. Alternatively, advertisers can select automatic targeting to allow Amazon to match their ads to relevant search terms and products. This advertising service is an important source of revenue for Amazon; it contributed US$14.1 billion (5.02%) to its *annual net sales* in 2019.[3]

As another example, Facebook launched its Dynamic Ads service to promote advertisers' products to people who expressed an interest in relevant keywords or similar products.[4] Dynamic Ads will automatically choose products from the catalog provided by the advertisers and display them to customers.

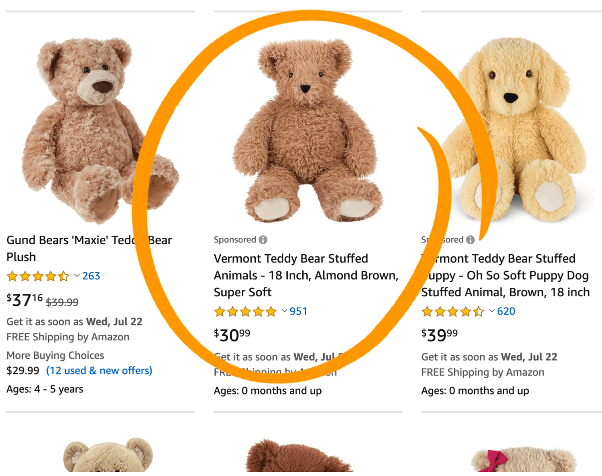


Figure 1 An Example of Sponsored Products on Amazon

Such large-scale e-commerce advertising platforms generally run thousands of advertising campaigns for different advertisers simultaneously. Each campaign is usually associated with (1) a budget that the advertiser wishes to spend as much as possible of during the campaign horizon and (2) a bid per CPC that dictates how much ad budget should be deducted upon each user click. The advertising platform dynamically allocates its ad spaces (i.e., customer impressions) to the ads whose campaigns

[2] For an explanation, see https://advertising.amazon.com/solutions/products/sponsored-products.

[3] See Amazon's 2019 financial report at https://www.sec.gov/ix?doc=/Archives/edgar/data/1018724/000101872420000004/amzn-20191231x10k.htm.

[4] See https://www.facebook.com/business/help/397103717129942?id=1913105122334058.

are active. As discussed above, an advertiser may require the platform to target his or her advertising campaign and ads to specific customer segments (specified by such features as location, age, and browsing, searching, and purchasing histories).

It is also not uncommon for advertisers and, thus, the platform to set click-through requirements for the ads (i.e., the minimal number of click-throughs during the entire ad campaign). For instance, Microsoft provides a Partner Incentive Cooperative Marketing Fund (Co-op) to subsidize its partners in whose website the number of click-throughs for Microsoft's ad is above 250 during the promotion events (Microsoft 2020). In addition, from a long-term perspective, the number of click-throughs for an ad substantially affects the long-term retention of the advertiser, which prompts the platform to devise the ad-allocation policy to secure a certain number of click-throughs for each ad (see, e.g., Ye et al. 2022).

To efficiently allocate its ad spaces, an online advertising platform faces the central operations problem of dynamically selecting a set of ads, which we refer to as an offer-set, displayed to each arriving customer in order to generate the highest total value throughout the planning horizon, subject to advertising budgets and click-through requirements for different ads.

Solving this problem presents a two-fold challenge. First, under customer-choice behavior, when an offer-set is displayed, the platform has to carefully balance the notorious trade-off in assortment optimization between expanding the offer-set to enlarge the market share and keeping it small enough to reduce cannibalization between different ads. Second, the click-through requirements, budget constraints, and advertisers' targeting rules altogether raise the difficulty of even searching for a feasible (but not necessarily optimal) ad allocation. In addition, we emphasize that the commonly used fluid-approximation (Fluid) approaches also face substantial computational challenges, because the number of decision variables increases exponentially with the cardinality of the offer-set.

In addition to optimizing total advertising value subject to budget and click-through-requirement constraints, the online platform also needs to address the fairness and discrimination concerns of its advertising/machine-learning algorithms. It is well-documented in the literature and in practice that a common source of algorithmic discrimination or bias in online advertising is that advertisers can target (or exclude) particular groups of users for their ads (e.g., Speicher et al. 2018).[5] The particular groups could be those classes harmed from counterfactual disparities, or legally protected characteristics, such as race and gender (e.g., Nilforoshan et al. 2022). Simple controls are insufficient to counter this issue, because the platform's underlying ad-allocation algorithm for optimizing certain business objectives, such as advertising revenue or advertisers' return on investment, may automatically skew ad delivery to certain user groups (see, e.g., Speicher et al. 2018, Lambrecht and Tucker 2019, Imana et al. 2021).

[5] See `https://www.reuters.com/technology/study-flags-gender-bias-facebooks-ads-tools-2021-04-09/`.

The main goal of this paper is to explore the e-commerce ad allocation of an advertising platform under customer choices and concerns for algorithmic fairness. Motivated by online e-commerce advertising practices, we seek to address the following key research question:

> *Taking into account algorithmic fairness, how should a platform dynamically personalize the ad offer-sets of each customer impression to maximize the fairness-adjusted value (FV) from advertising throughout a planning horizon in the presence of budget constraints and click-through requirements for different ads?*

1.1. Main Contributions

In this paper, we present a general stochastic-program model to study the complex dynamic ad-allocation problem and propose a scalable algorithm that solves this problem to (asymptotic) optimality. Our key contribution is proposing a novel *two-stage target-debt* (TTD) framework that yields a simple, computationally efficient and asymptotically optimal policy for addressing this otherwise intractable online ad-allocation problem for an e-commerce platform. This TTD framework carefully combines three ideas: click-through target optimization, compact reformulation, and a debt-weighted offer-set algorithm.

Click-through target optimization. In the first-stage, we approximate the original problem as an (auxiliary) click-through target optimization, in which the platform decides the click-through goal (targets) for each ad-customer pair to maximize the expected FV. Such a reduction to a deterministic convex program is not only tractable, but also provides a new upper bound for the original (intractable) stochastic program for ad-allocation optimization.

Compact reformulation. We characterize the necessary and sufficient condition under which the click-through targets are feasible. The characterization is surprisingly simple for most commonly used choice models such as multinomial logit (MNL), independent, and generalized attraction choices. In these cases, we can efficiently solve the click-through target optimization using the first-stage convex program, because the scale of our reformulation increases linearly (rather than exponentially) with the number of ads.

Debt-weighted offer-set algorithm. In the second-stage, an ad offer-set is displayed to each user upon his or her arrival in order to satisfy the optimal click-through targets set in the first stage. We propose a simple and effective algorithm, referred to as the *debt-weighted offer-set* (DWO) policy. This policy dynamically assigns a "debt" to each click-through target that measures the difference between the realized total click-throughs and the endogenous target set in the first-stage convex program. Then, a standard offer-set/assortment optimization problem is solved to maximize a debt-weighted value function upon the arrival of each customer.

We prove that our TTD framework yields an asymptotic optimal policy (i.e., the DWO algorithm initialized with the optimal click-through targets) for the online ad-allocation problem. Consistent with most existing debt-weighted algorithms (e.g., Zhong et al. 2017, Jiang et al. 2023), our DWO algorithm satisfies the feasibility and approchability of click-through targets. Our refined analysis also provides the optimality guarantee of the TTD framework. Leveraging the *exact-approachability* of the DWO algorithm to *exactly* meet the (feasible) target set in the first stage, we establish new intrinsic connections and asymptotic equivalence of the original ad-allocation problem, the first-stage convex program, and an auxiliary Fluid convex program, implying that the theoretical upper bound of FV characterized by the first-stage convex program can be achieved. If the first-stage click-through target vector is only feasible, but not optimal, the associated DWO policy will not incur any additional optimality loss on top of that from the suboptimal target in the asymptotic regime.

The TTD framework we propose in this paper is computationally scalable if customer choices follow commonly adopted models such as MNL, independent, and generalized attraction choices. Through numerical experiments, we demonstrate that our algorithm outperforms the commonly adopted Fluid-based algorithms (e.g., Liu and Van Ryzin 2008, Jasin and Kumar 2012, Bumpensanti and Wang 2020, Balseiro et al. 2023) in terms of performance, robustness, and computational efficiency for most problem instances. Our numerical experiments also demonstrate that the proposed policy depletes the budgets much more smoothly than the benchmarks over the entire planning horizon. This highlights the practical applicability of our approach, because smooth budget depletion is a desirable property for the real-world online advertising business. In addition, our approach helps substantially improve the algorithmic fairness of ad allocation for an online e-commerce platform without compromising its efficiency much, achieving high FV with low variance. Hence, our computationally light TTD framework well handles the notorious bias and discrimination issues in online ad allocation.

In summary, the key takeaway from this paper is that the proposed TTD framework, which induces a two-stage stochastic program reformulation combining click-through target optimization and the debt-weighted off-set algorithm, can efficiently address the ad-allocation problem for e-commerce platforms to improve the FV from advertising. Our approach is simple, efficient, and scalable, with a provable optimality guarantee and strong numerical performances.

The rest of this paper is organized as follows. We review the related literature in Section 2. We introduce the model in Section 3, and we propose the two-stage stochastic program reformulation in Section 4. We study the optimal ad-allocation policy in Section 5, and we present the numerical studies in Section 6. Section 7 concludes. All proofs are relegated to the Online Appendices. Throughout this paper, we use **boldface** to represent vectors and matrices, and $\boldsymbol{a} \geq \boldsymbol{b}$ to represent $a_j \geq b_j$ for each j. For notational conciseness, we do not differentiate a random variable and its realization whenever there is no ambiguity.

2. Literature Review

This paper proposes a general framework and an efficient algorithm to study optimal online ad allocation for an e-commerce platform under algorithmic fairness concerns. Our paper is primarily related to four streams of research in the literature: (a) ad allocation for online advertising platforms, (b) algorithmic discrimination/bias in online advertising, (c) resource allocation with individualized service-level constraints, and (d) (dynamic) personalized assortment optimization. Papers in the literature generally focus on one or more of the four topics above, whereas our work contributes to all four streams of literature jointly.

Ad allocation is a central challenge for online advertising platforms. Scheduling advertising displays on websites has been widely studied in the literature. For example, Nakamura and Abe (2005) propose an ad-targeting approach based on linear programming that achieves high click-through rates by optimizing ad-display probabilities. Furthermore, a queue is adopted that holds the selected ads to avoid displaying the same ad more than once on one page and to reduce the display probabilities with small margins. Yang et al. (2012) combine an ad-inventory allocation problem with the multi-objective of revenue and fairness. For maximizing the reach of customers and minimizing the variance of the outcome simultaneously in targeted advertising, Turner (2012) formulates an-planning problem with a quadratic objective to spread ads across all targeted customer types. Balseiro et al. (2014) formalize an ad-exchange problem as a multi-objective stochastic control problem considering both the revenue from exchange and click-through rates, and they derive an efficient policy for online ad allocation with uncertainty. Also for dealing with uncertainty, Shen et al. (2021a) propose an integrated planning model with a distributionally robust chance-constrained program in online ad allocation. Hojjat et al. (2017) consider a new contract to allow advertisers to specify the number of unique individuals who should see their ad and the minimum number of times each individual should be exposed. They also introduce a new mechanism for ad serving that “pre-generates” an explicit sequence of ads for each user to see over time. Shen et al. (2021b) deal with customers’ ad-clicking behavior by an arbitrary-point-inflated Poisson regression model, and they solve a mixed-integer nonlinear programming model for optimal ad allocation. We refer interested readers to Choi et al. (2020) for a comprehensive review of this literature. The key modeling difference of our paper from this literature is that, using choice models, we clearly model the click-through behaviors of a customer in the presence of multiple ads displayed to him or her simultaneously.

As mentioned earlier, algorithmic discrimination/bias in online advertising has received increased scrutiny in recent literature. Several works find evidence of algorithmic discrimination/bias based on race or gender in practical online advertising platforms. Speicher et al. (2018) show that simply disallowing the use of some sensitive features, such as race and gender, in targeted advertising is not sufficient to remove algorithmic bias. By exploring the data from a field test, Lambrecht and Tucker

(2019) find that an algorithm delivered fewer job-opportunity ads to women than to men because it was more expensive to deliver the ads to young women. Though the algorithm optimizes ad delivery to achieve cost-effectiveness in a gender-neutral way, its intrinsic mechanism autonomously leads to bias and discrimination. Ali et al. (2019) find that, despite the neutral targeting parameters, both the budget and the content of the ad significantly contribute to Facebook's ad-delivery bias along gender and racial lines for ads on employment and housing opportunities. Imana et al. (2021) develop a new methodology for identifying discrimination due to protected categories such as gender and race in the delivery of job advertisements, distinguish this discrimination from legal bias due to differences in qualification among users, and confirm this discrimination by gender in ad delivery on Facebook.

To mitigate the algorithmic discrimination/bias in online advertising, Celis et al. (2019) present a constrained ad-auction framework that maximizes the platform revenue subject to the condition that the number of users seeing an ad is distributed appropriately across sensitive user types such as gender and race. They solve a nonconvex optimization problem to obtain the optimal auction mechanism for a large class of fairness constraints. Lejeune and Turner (2019) derive a Gini-index-based metric to measure how well dispersed the impressions are allocated across audience segments, and they formulate an optimization problem to maximize the spread of impressions across targeted audience segments while minimizing demand shortfalls.

Balseiro et al. (2021) use a nonlinear regularizer as a fairness measure, and they design an online resource-allocation algorithm to maximize the weighted objective of efficiency and fairness subject to the resource constraints. Considering a single server system where individual customers are differentiated in their service requirements and waiting costs, Mulvany and Randhawa (2021) study a fairness-efficiency trade-off, which leads to societal inequities in their operational metrics. Bateni et al. (2022) adopt a weighted proportional fairness metric under the setting that a platform dynamically allocates to budgeted buyers a collection of goods that arrive to the platform online. In this paper, we offer a new slant on mitigating algorithmic discrimination/bias in online advertising by considering different fairness metrics in the above literature and *disparate impact*, which identifies unintentional bias of an algorithm (see, e.g., Feldman et al. 2015).

The resource-allocation problem of meeting service-target constraints in the face of uncertain demand has been extensively studied in the inventory literature (see, e.g., Eppen 1979, Swaminathan and Srinivasan 1999, Alptekinoğlu et al. 2013). Leveraging Blackwell's approachability theorem, Zhong et al. (2017) characterize the optimal safety-stock level with individual type-II service-level constraints. Lyu et al. (2019) and Lyu et al. (2022), respectively, extend both the approach and the results to the context of type-I service-level constraints and process flexibility. Utilizing a semi-infinite linear program formulation, Jiang et al. (2023) generalize and unify models in this literature and propose a simple randomized rationing policy to meet general service-level constraints, including type-I and type-II constraints, and beyond. Ma et al. (2020) consider an online-matching problem with concerns of agent-group fairness; they define two different service-level objectives—instead of maximizing the number of matches—as the metrics for long-run and short-run fairness, and they show the competitive ratios of their algorithms. We contribute to this strand of the literature by generalizing the concept of service-level constraints to incorporate customer choice uncertainty and ad allocation through assortment planning.

We also propose a debt-weighted offer-set algorithm and demonstrate its optimality for meeting the endogenous service targets and for generating the total payoff for the platform. Our debt-weighted offer-set policy and other existing debt-weighted policies are closely related to the well-known max-weight policy proposed in the queueing literature (e.g, Tassiulas and Ephremides 1990, McKeown et al. 1999, Stolyar 2004), which is also commonly used in the resource-allocation and capacity-planning literature. Dai and Lin (2005) propose a max-weight policy for dynamically allocating service capacities in a stochastic processing network, and they prove that the policy can achieve maximum throughput predicted by a deterministic model under a mild assumption.

By applying a max-weight policy with good performance, Shi et al. (2019) analyze the design of sparse flexibility structures in a multiperiod make-to-order production system. Xu and Zhong (2020) formulate a generalized version of max-weight policy to study the impact of information constraints and memories on dynamic resource allocation. We refer interested readers to the above works for more details on max-weight policy. In this paper, we adopt target-based weight updating, which utilizes both the targets optimized in the first-stage program and the realized click-throughs—rather than just the current states, as shown in this literature—to adjust the weights in each period.

Over the past ten years, online e-commerce platforms have typically provided numerous products for customers to choose from (Feldman et al. 2022). Manufacturing firms have also expanded their product lines due to business trends (e.g., fast fashion; Caro et al. 2014) or technology revolution (e.g., 3D printing; Dong et al. 2022). The ever-expanding product pool makes personalized assortments more attractive. Therefore, personalized-assortment optimization has also been receiving increased attention in the literature.

Leveraging the competitive-ratio framework, Golrezaei et al. (2014) propose inventory-balancing algorithms that guarantee the worst-case revenue performance without any forecast of the customer type distribution. Bernstein et al. (2019) combine dynamic assortment planning, demand learning, and customer-type clustering in a Bayesian framework, and they propose a prescriptive assortment-personalization approach for online retailing. Using re-solving heuristics, Jasin and Kumar (2012) study the dynamic, personalized-product-offering problem in a network revenue-management framework, and they show that the proposed heuristics achieve a constant optimality loss. Bumpensanti and Wang (2020) design a new, infrequent re-solving heuristic that re-solves the deterministic linear program less frequently at the beginning of the planning horizon and more frequently at the end. This infrequent re-solving heuristic achieves a revenue loss uniformly bounded by a constant that's independent of the time-horizon length and resource capacities.

Meanwhile, Kallus and Udell (2020) consider a dynamic assortment-personalization problem with a large number of items and customer types as a discrete contextual-bandit problem, and they propose a structural approach with efficient optimization algorithms. Chen et al. (2023) formalize a new checkout-recommender system at Walmart's online grocery store as an online assortment-optimization problem with limited inventory, and they propose an inventory-protection algorithm with a bounded competitive ratio. Gallego et al. (2016) study a general personalized resource-allocation model with customer choices. Adopting the column-generation approach to solve the choice-based linear program, the authors introduce algorithms with theoretical performance guarantees. Considering the uncertainty in estimating the MNL choice model, Cheung and Simchi-Levi (2017) propose a Thompson-sampling-based policy to estimate the latent parameters by offering a personalized assortment, and they demonstrate its near optimality. We contribute to this literature by proposing a new, two-stage stochastic program framework (i.e., TTD) to study the ad offer-set optimization problem. Moreover, we design a novel DWO policy that proves to be asymptotically optimal and generates values with lower variance than the Fluid benchmarks commonly adopted in the literature.

3. Model

3.1. Model Setup

The platform and its customers. We consider an e-commerce platform such as Amazon or Facebook Marketplace that matches its user traffic with product advertisements. The advertisements thereof are usually labeled as *Sponsored Products*, as shown in Figure 1. Our model, however, can be straightforwardly applied to the setting of product recommendation on an e-commerce platform. Throughout the planning horizon, there are T customer impressions (also called users or viewers) arriving at the platform sequentially. So we say customer t arrives in period t. Without loss of generality, we assume T is deterministic and known to the platform. Customers are segmented into m types based on their demographic information (e.g., age, gender, location) and behavior on the platform (e.g., average spending per year, product preferences, average time spent on the platform per year). We denote $\mathcal{M} := \{1, 2, ..., m\}$ as the set of all customer segments. For each customer t, his or her type $j(t)$ is *i.i.d.* and follows a discrete distribution on $\mathcal{M}$, with $\mathbb{P}(j(t) = j) = p^j > 0$, where $j \in \mathcal{M}$ and $\sum_{j\in\mathcal{M}} p^j = 1$.

Advertisements. At the beginning of the horizon, advertisers launch a set of ad campaigns, which we denote as $\mathcal{N} := \{1, 2, ..., n\}$. For each ad campaign $i \in \mathcal{N}$, its advertiser sets $B_i > 0$ as the total budget and $b_i > 0$ as the bid price, which is a proxy for the ex-post price per click paid by the advertiser to the platform, with the exact auction setting abstracted away from our model.[6] Specifically, B_i is the maximum advertising fee the advertiser will pay the platform throughout the ad campaign's life, and the budget will be depleted by b_i upon each click by a customer. That is, the platform adopts the CPC mechanism, which is commonly used in online advertising. Furthermore, the ads are placed in specific slots exclusively allocated to advertising.

Offer-sets. Upon the arrival of user t, the platform observes its type $j(t)$ and decides a (possibly randomized) set of ads/sponsored products displayed to this user (which we call an *offer-set*), denoted by $S(t) \in \mathfrak{S}^{j(t)} \subset 2^{\mathcal{N}}$, where $\mathfrak{S}^j$ is the set of all feasible offer-sets to type-j customers, and $2^{\mathcal{N}}$ is the power set of $\mathcal{N}$. Throughout our analysis, we make the following assumption on the structure of $\mathfrak{S}^j$.

Assumption 1. *If $S \in \mathfrak{S}^j$, then for any subset $S' \subset S$, we also have $S' \in \mathfrak{S}^j$.*

Assumption 1 implies that $\emptyset \in \mathfrak{S}^j$, i.e., the platform may not display any ad to a customer of type j. We may impose additional structural constraints on $\mathfrak{S}^j$. Of particular importance is the cardinality constraint (i.e., the total number of ads displayed to customers cannot exceed K; see, e.g., Rusmevichientong et al. 2010, Wang 2012, Davis et al. 2013), which is prevalent in the online advertising setting where the platform cannot allocate more ads to a customer than the number of available ad-impression slots. We will demonstrate how to handle this cardinality constraint in our theoretical and numerical analyses below.

Click-throughs. For user t, if an offer-set $S(t)$ is displayed (see Figure 1), she may or may not click some ads in the set $S(t)$. We denote $y_i^j(t)$ as the indicator of one click received by ad i from a type-j customer in time t. Therefore, $y_i^j(t) = 1$ only if $j(t) = j$, $i \in S(t)$, and customer t clicks ad i. Otherwise, $y_i^j(t) = 0$. We denote the click-through matrix in time t as $\boldsymbol{y}(t) := (y_i^j(t) : i \in \mathcal{N}, j \in \mathcal{M})$. We do not specify any structure of the customer click-through behavior. For each customer type j, each offer-set S, and each ad $i \in S$, we denote the expected click-throughs of ad i in an offer-set S from a type-j customer as $\phi_i^j(S) := \mathbb{E}_{\mathsf{D}_y}[y_i^j(t) | j(t) = j, S(t) = S]$, where D_y is the click-through distribution. We denote $\mathsf{D}_{(j,y)}$ as the joint customer type and click-through distribution.

We assume customers exhibit (conditionally) independent and stationary click-through behaviors. Specifically, conditioned on the realized offer-sets $\{S(\tau) : 1 \le \tau \le T\}$, the click-throughs, $\boldsymbol{y}(t)$'s, are independent across time t. Furthermore, conditioned on the same offer-set, $\boldsymbol{y}(t)$'s are identically

[6] For ease of exposition, we consider the same bid price of advertiser i across all customer types (i.e., b_i). In practice, an advertiser may set different bids for different targeted groups. If we allow viewer-type-dependent bid price b_i^j, our analysis and results will not be affected.

distributed with respect to the time t, i.e., for any set of click-through outcomes $\mathcal{Y}$, any realized customer type j, any offer-set S, and any $t \neq \tau$, $\mathbb{P}[\boldsymbol{y}(t) \in \mathcal{Y} | j(t) = j, S(t) = S] = \mathbb{P}[\boldsymbol{y}(\tau) \in \mathcal{Y} | j(\tau) = j, S(\tau) = S]$. By stationarity, the function $\phi_i^j(\cdot)$ is independent of time t. If $i \notin S$, $\phi_i^j(S) = 0$ by definition.

Ad targeting and click-through requirement. From the advertisers' perspective, they target ads to the relevant customer segments based on their demographic information, past behavioral patterns, and potential interests (see, e.g., Choi et al. 2020). Furthermore, consistent with advertising practices (e.g., Microsoft 2020), the advertiser may require that ad i receives at least $\eta_i^{\mathcal{C}}$ click-throughs throughout the planning horizon for targeting a set of customer types $\mathcal{C} \in \mathfrak{K}_i$, where $\mathfrak{K}_i \subset 2^{\mathcal{M}}$ denote the set of all customer segment subsets $\mathcal{C}$ on which the advertiser sets a positive click-through requirement $\eta_i^{\mathcal{C}} > 0$ of ad i. Mathematically, the targeting and click-through requirement of ad i can be formalized as $\sum_{t=1}^{T} \sum_{j \in \mathcal{C}} y_i^j(t) \geq \eta_i^{\mathcal{C}}$ for any $i \in \mathcal{N}$ and $\mathcal{C} \in \mathfrak{K}_i$. In practice, $\mathfrak{K}_i$ often only contains either $\mathcal{M}$ (i.e., the requirement of total click-throughs) or some non-overlapping subsets of $\mathcal{M}$. For example, Pampers may additionally ask the platform to target its diaper ads to new parents and babysitters. Note that, due to the randomness in the customer types and choices, the above targeting and click-through requirements may not be satisfied with probability 1. Hence, we model the click-through requirements in the expected sense, i.e., $\mathbb{E}\left[\sum_{t=1}^{T} \sum_{j \in \mathcal{C}} y_i^j(t)\right] \geq \eta_i^{\mathcal{C}}$, which is a common modeling approach in the literature on resource allocation to meet service target constraints (e.g., Zhong et al. 2017, Jiang et al. 2023). We will also show in Theorem 2 that our proposed policy can indeed satisfy the click-through requirements almost surely in the asymptotic regime where the problem size scales to infinity. In Appendix B, we illustrate how to incorporate the click-through requirements as soft-constraints, i.e., by adding a penalty term $-\left(\eta_i^{\mathcal{C}} - \sum_{t=1}^{T} \sum_{j \in \mathcal{C}} y_i^j(t)\right)^+$ for each $i \in \mathcal{N}$ and $\mathcal{C} \in \mathfrak{K}_i$ into the objective function.

Moreover, large-scale online advertising platforms (e.g., Facebook and Google) have recently tightened their controls to prevent advertisers from excluding some user segments in their target in order to reduce lawsuits and regulatory probes into discrimination. Thus, in most cases, the total number of click-through requirement constraints $|\mathfrak{K}_i|$ is at most *linear* (instead of exponential) in the total number of customer segments m, making our model and solution approach scalable. In many scenarios, the advertising contract specifies the minimal click-through requirement. For instance, Microsoft (as an advertiser) requires its partners (i.e., the advertising platforms where Microsoft runs its advertising campaigns) to earn at least 250 click-throughs during one ad campaign to be qualified to receive the support through its Co-op.

Advertising value and fairness. The total value of online advertising generated throughout the planning horizon depends on matching the n ads with T customers. Specifically, each click of ad i by a type-j customer generates a value of $r_i^j \geq 0$, which is allowed to be both ad- and customer-type-specific. The interpretation of r_i^j, which can be quite general, includes the following scenarios as special cases. For the case where the value is the total advertising revenue of the platform, $r_i^j = b_i$ for each $i \in \mathcal{N}$. For the case where the value is the total advertising return of the advertisers (see, e.g., Hao et al. 2020), r_i^j is interpreted as the value of one click-through for ad i by a type-j customer to its advertiser. Therefore, the total value of online advertising is given by:

$$\sum_{t=1}^{T} \sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{M}} r_i^j y_i^j(t)$$

A salient feature of our model is that, in addition to total advertising value, the platform may also be concerned about the *fairness* of the system. For example, the recent advocacy on machine learning/algorithmic fairness postulates that customers who are considered minorities should have sufficient click-throughs/conversions in a recommender/advertising system; otherwise, their needs cannot be well taken care of due to data sparsity (see, e.g., Lambrecht and Tucker 2019). In a similar vein, advertisers generally prefer receiving impressions that are evenly spread across their targeted customer types (see, e.g., Lejeune and Turner 2019). To account for such algorithmic fairness, we introduce a general fairness metric $F(\cdot): \mathbb{R}^{nm} \mapsto \mathbb{R}$, which is a function of the per-customer-impression click-through matrix: $\bar{\boldsymbol{y}} := (\bar{y}_i^j : i \in \mathcal{N}, j \in \mathcal{M})$, where $\bar{y}_i^j := \frac{1}{T}\sum_{t=1}^{T} y_i^j(t)$ is the per-customer-impression click-through of ad i by type-j customers. Throughout our analysis, we make the following assumption on the concavity of $F(\cdot)$.

Assumption 2. *The fairness metric $F(\cdot): \mathbb{R}^{nm} \mapsto \mathbb{R}$ is a concave function of the per-customer-impression click-through matrix $\bar{\boldsymbol{y}}$.*

Because there is no universally accepted quantitative definition of fairness and equality, the fairness metric $F(\cdot)$ may be one of many forms in different contexts, as detailed in Appendix C. For example, $F(\cdot)$ can be modeled as *allocative fairness* measures (e.g., max-min fairness, Gini mean difference (GMD) fairness) to keep the advertisers from exclusively targeting a small subset of user groups, or $F(\cdot)$ can be modeled as *disparate impact* measures (see, e.g., Feldman et al. 2015, Barocas and Selbst 2016) to reduce algorithmic discrimination. Therefore, we measure the fairness in a *per-customer-impression* sense and evaluate the *per-customer-impression* FV from advertising as:

$$\frac{1}{T} \sum_{t=1}^{T} \sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{M}} r_i^j y_i^j(t) + \lambda F(\bar{\boldsymbol{y}})$$

where $\lambda \geq 0$ parameterizes the trade-off between efficiency and fairness. The smaller (resp. larger) the λ, the higher weight the platform puts on efficiency (resp. fairness). In the extreme case where $\lambda \to 0$

(resp. $\lambda \to +\infty$), the system is purely efficiency-driven (resp. fairness-driven). We also remark that the concavity of $F(\cdot)$ is a mild assumption, which can be satisfied by most of the commonly adopted fairness metrics in the literature, such as max-min fairness, α-fairness, GMD fairness, disparate impact, and so on (see, e.g., Balseiro et al. 2021).

3.2. Stochastic Program Formulation

We consider the (randomized) policies Π. First, we define the realized history until time t as $\mathcal{H}_{t-1} := \{(j(\tau), S(\tau), \boldsymbol{y}(\tau)) : 1 \le \tau \le t-1\}$. By convention, $\mathcal{H}_0 = \emptyset$. In time t, π maps the realized customer type $j(t)$ and the realized history $\mathcal{H}_{t-1}$ to a distribution on all feasible offer-sets to a type-$j(t)$ customer, $\mathfrak{S}^{j(t)}$, i.e., $S(t) = \pi(j(t), \mathcal{H}_{t-1})$. Note that deterministic policies are special cases of Π, which map $(j(t), \mathcal{H}_{t-1})$ to a deterministic offer-set in $\mathfrak{S}^{j(t)}$. Sometimes, it is useful to spell out the dependence of the click-through outcomes in time t, $\boldsymbol{y}(t)$, on the history $\mathcal{H}_{t-1}$ and the policy π. We use $y_i^j(t|\pi)$ as the number of click-throughs for ad i by a type-j customer in time t, given that the history is $\mathcal{H}_{t-1}$ and the offer-set displayed to a type-$j(t)$ customer is $S(t) = \pi(j(t), \mathcal{H}_{t-1})$. Likewise, we define $\bar{y}_i^j(\pi) := \frac{1}{T}\sum_{t=1}^{T} y_i^j(t|\pi)$ as the per-customer click-throughs of ad i by type-j customers in the entire horizon under policy π. We denote $\bar{\boldsymbol{y}}(\pi) := (\bar{y}_i^j(\pi) : i \in \mathcal{N}, j \in \mathcal{M})$ as the per-customer-impression click-through matrix under policy π.

Of particular importance are the (randomized) *static* policies, the set of which we denote as $\Pi_{\mathsf{static}} \subset \Pi$. Specifically, if $\pi \in \Pi_{\mathsf{static}}$, the offer-set displayed in time t, $S(t) = \pi(j(t), \mathcal{H}_{t-1})$ is independent of (i) time t and (ii) the history $\mathcal{H}_{t-1}$, solely conditioned on the realized customer type $j(t)$. Hence, for $\pi \in \Pi_{\mathsf{static}}$, we can drop the history $\mathcal{H}_{t-1}$ to write $\pi(j(t), \mathcal{H}_{t-1})$ as $\pi(j(t))$. The word "static" also refers to that the distribution of $S(t)$ is stationary with respect to time t for each customer type j. Therefore, we denote π_{static} as a static policy to distinguish it from an arbitrary policy $\pi \in \Pi$.

We are now ready to formulate the platform's ad-allocation problem as a multiperiod stochastic program. Specifically, the platform seeks to optimize the total expected FV of online advertising throughout the planning horizon:

$$
\begin{aligned}
&\max_{\pi \in \Pi} \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^j y_i^j(t|\pi) + \lambda F(\bar{\boldsymbol{y}}(\pi))\right] \\
&\text{s.t. } \frac{1}{T}\sum_{t=1}^{T}\sum_{j\in\mathcal{M}} b_i y_i^j(t|\pi) \le \frac{B_i}{T}, \ \textit{almost surely} \text{ for each } i \in \mathcal{N}, \\
&\qquad \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}\sum_{j\in\mathcal{C}} y_i^j(t|\pi)\right] \ge \frac{\eta_i^{\mathcal{C}}}{T}, \text{ for each } i \in \mathcal{N} \text{ and } \mathcal{C} \in \mathfrak{K}_i
\end{aligned}
\tag{$\mathcal{OP}$}
$$

where the first term in the objective is the total per-customer-impression value from advertising, which we call the *efficiency* of policy π denoted by $\mathcal{E}(\pi) := \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^j y_i^j(t|\pi)\right]$, the second term is the *fairness* of policy π denoted by $\lambda \cdot \mathcal{F}(\pi) := \lambda \cdot \mathbb{E}[F(\bar{\boldsymbol{y}}(\pi))]$, and all the expectations including the following ones are taken with respect to policy π and $\mathsf{D}_{(j,y)}$ unless otherwise stated. Hence, the total FV under policy π is given by $\mathcal{V}(\pi) := \mathcal{E}(\pi) + \lambda \cdot \mathcal{F}(\pi)$. We also remark that the first constraint of $(\mathcal{OP})$ refers to the budget constraint of each ad and the second refers to the click-through requirement for each ad with respect to different sets of customer types in the expected sense. We denote the optimal FV of $(\mathcal{OP})$ as $\mathcal{V}^* = \limsup_{\pi \in \Pi} \mathcal{V}(\pi)$ and the optimal policy (if it exists) as $\pi^* = \arg\max_{\pi \in \Pi} \mathcal{V}(\pi)$.

Roadmap to solve $(\mathcal{OP})$. For the rest of this paper, our goal is to design a two-stage ad-allocation framework (i.e., TTD) to find a policy π that achieves the optimal FV, $\mathcal{V}^*$, while satisfying the budget- and click-through-requirement constraints. As detailed below, our TTD framework to solve $(\mathcal{OP})$ can be decomposed into the following two stages:

- **First-stage click-through target optimization,** where we solve a deterministic (but *nonequivalent*) convex program to obtain the optimal click-through target of the ads and customer types (Section 4).
- **Second-stage offer-set allocation,** where we adaptively decide the offer-set displayed to each customer based on how far away the click-throughs are from the optimal targets obtained in the first-stage by a debt-weighted offer-set policy (Section 5.1).

We demonstrate that the TTD framework is asymptotically optimal (Section 5.2) and enjoys impressive performance in the nonasymptotic regime compared with the commonly adopted benchmarks in the existing literature (Section 6).

4. Reformulation and Feasibility Conditions

In this section, we propose a novel reformulation of the original intractable ad-allocation program $(\mathcal{OP})$ to maximize the expected FV as a much simpler, two-stage convex optimization. The core of our reformulation is to introduce the auxiliary click-through targets of the ads by different customer types, then to design an online debt-based ad-allocation algorithm to achieve these targets. We also emphasize that although the reformulated convex optimization is *not* necessarily equivalent to $(\mathcal{OP})$, the proposed algorithm is indeed *provably optimal* for the original ad allocation in the asymptotic regime, where the problem size scales up to infinity.

4.1. Problem Reformulation With Click-Through Targets

To solve the dynamic ad-allocation problem $(\mathcal{OP})$, a commonly adopted approach in the literature is to consider a fluid approximation of this problem and solve the Fluid problem by linear programming (choice-based linear programming (CDLP); see, e.g., Liu and Van Ryzin 2008, in our case, the Fluid problem is a convex program) and linear programming-resolving (LP-resolving) heuristics

(LP-resolving; see, e.g., Jasin and Kumar 2012, Bumpensanti and Wang 2020). In particular, we refer interested readers to Balseiro et al. (2023) where they introduce a unifying model for a general class of dynamic resource allocation problem and analyze the performance of a Fluid-based heuristic with resolving at each period. The Fluid-based formulation of ($\mathcal{OP}$) is provided by ($\mathcal{OP}_{\mathsf{Fluid}}$) in Section 5.3 as an auxiliary problem to demonstrate the optimality of our proposed algorithm. With the cardinality constraint on the feasible offer-sets, one difficulty using (the analogs of) CDLP or LP-resolving heuristics is that the number of variables (i.e., the probability of each offer-set for all customer types) quickly explodes as the number of products increases, even when the choice model is as simple as the independent choice or MNL model.

To tackle the aforementioned challenges of the standard Fluid approach, we develop a novel reformulation of ($\mathcal{OP}$) that transforms the original problem as a two-stage convex optimization by introducing click-through targets associated with each ad-customer pair as auxiliary decision variables. Such reformulation also proves useful to design our asymptotically optimal ad-allocation algorithm. Specifically, we define $\boldsymbol{\alpha} := (\alpha_i^j, i \in \mathcal{N}, j \in \mathcal{M}) \in \mathbb{R}_+^{nm}$, where α_i^j refers to the (virtual) target for the per-customer-impression number of click-throughs for ad i by type-j customers. So the platform operationalizes its ad-allocation algorithm such that the total number of click-throughs for ad i by type-j customers exceeds $T\alpha_i^j$. We define the concave per-customer-impression FV associated with click-through target vector $\boldsymbol{\alpha}$ as

$$\mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) := \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^j \alpha_i^j + \lambda F(\boldsymbol{\alpha}) \tag{1}$$

where the first term captures efficiency, the second captures fairness with respect to the click-through target vector $\boldsymbol{\alpha}$, and the subscript "CT" stands for *click-through target.* We transform ($\mathcal{OP}$) into the following (nonequivalent) optimization problem:

$$\begin{aligned} \max_{\pi\in\Pi,\boldsymbol{\alpha}\geq\mathbf{0}} \; & \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) \\ \text{s.t. } & \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} y_i^j(t|\pi)\right] \geq \alpha_i^j, \quad \text{for each } i\in\mathcal{N}, j\in\mathcal{M}, \\ & b_i \sum_{j\in\mathcal{M}} \alpha_i^j \leq \frac{B_i}{T}, \text{ for each } i\in\mathcal{N}, \\ & \sum_{j\in\mathcal{C}} \alpha_i^j \geq \frac{\eta_i^{\mathcal{C}}}{T}, \text{ for each } i\in\mathcal{N} \text{ and } \mathcal{C}\in\mathfrak{K}_i \end{aligned} \tag{2}$$

Comparing (2) with ($\mathcal{OP}$) reveals that our reformulation relaxes the sample-path-based objective function and constraints in the original problem with their counterparts characterized by the per-customer-impression click-through target vector $\boldsymbol{\alpha}$. To ensure that the reformulation is close enough to the original problem and that the click-through targets are achievable, we introduce an additional constraint: the expected click-throughs per customer should meet the click-through targets, as specified by the first constraint of (2). We emphasize that (2) is *not* equivalent to the ($\mathcal{OP}$) in general. However, we show in Section 5 that in the asymptotic regime where the problem size scales up to infinity, there is an algorithm based on the solution to (2) achieving the optimal FV of ($\mathcal{OP}$). In this sense, our reformulation is *asymptotically equivalent* to the original dynamic ad-allocation problem.

It is still challenging to characterize when the *expected* click-through target constraint in (2), i.e., $\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} y_i^j(t|\pi)\right] \geq \alpha_i^j$, can be satisfied. Therefore, we further modify (2) by replacing this constraint with one for the expected number of click-throughs under static policies Π_{static}. Specifically, we replace the first constraint of (2) with

$$\mathbb{E}[y_i^j(t|\pi_{\mathsf{static}})] \geq \alpha_i^j, \text{ for each } i \in \mathcal{N}, j \in \mathcal{M} \tag{3}$$

One should note that the expected click-through target constraint (3) is independent of time t. Hence, we reformulate the ad-allocation problem by further modifying the click-through target constraint as follows:

$$\begin{aligned} \max_{\pi \in \Pi_{\mathsf{static}}, \boldsymbol{\alpha} \geq \mathbf{0}} & \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) \\ \text{s.t. } & \mathbb{E}[y_i^j(t|\pi)] \geq \alpha_i^j, \text{ for each } i \in \mathcal{N}, j \in \mathcal{M}, \\ & b_i \sum_{j \in \mathcal{M}} \alpha_i^j \leq \frac{B_i}{T}, \text{ for each } i \in \mathcal{N}, \\ & \sum_{j \in \mathcal{C}} \alpha_i^j \geq \frac{\eta_i^{\mathcal{C}}}{T}, \text{ for each } i \in \mathcal{N} \text{ and } \mathcal{C} \in \mathfrak{K}_i \end{aligned} \tag{2$\mathcal{SSP}$}$$

It is clear from (2$\mathcal{SSP}$) that the original problem is reformulated as a *two-stage stochastic program*. In the first stage, the platform selects the click-through targets $\boldsymbol{\alpha}$ to maximize a variant of the FV, $\mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha})$; in the second stage, it selects a static policy π_{static} to meet the first-stage click-through targets (3). Compared with (2), (2$\mathcal{SSP}$) has a more stringent constraint (3) given that $\Pi_{\mathsf{static}} \subset \Pi$. In other words, (2$\mathcal{SSP}$) provides a lower bound of (2). However, we demonstrate their equivalence in the following lemma.

Lemma 1. *A click-through target vector $\boldsymbol{\alpha}$ is feasible to (2$\mathcal{SSP}$) if and only if it is feasible to (2). Furthermore, any optimal click-through target vector $\boldsymbol{\alpha}$ of (2$\mathcal{SSP}$) is also optimal for (2), and vice versa.*

Lemma 1 suggests that to find the optimal click-through targets of (2), it suffices to solve (2$\mathcal{SSP}$). Indeed, we show in Section 5 that there is an optimal randomized static policy for (2$\mathcal{SSP}$), i.e., the Fluid policy, achieving the optimal FV of ($\mathcal{OP}$) in the asymptotic regime where the problem size scales to infinity. This helps us establish the asymptotic equivalence of these two reformulations and the original ad-allocation problem.

We call a click-through target vector $\boldsymbol{\alpha}$ *single-period feasible* if there exists a static policy $\pi_{\mathsf{static}} \in \Pi_{\mathsf{static}}$ such that (3) holds. The single-period feasibility condition for a click-through target vector $\boldsymbol{\alpha}$ is central to the design and analysis of our algorithm to solving both ($2\mathcal{SSP}$) and, eventually, ($\mathcal{OP}$). The rest of this section will be devoted to characterizing the necessary and sufficient condition for an $\boldsymbol{\alpha}$ to be single-period feasible.

Sometimes it is more convenient to rewrite this expected click-through target condition (3) as a periodic-review infinite-horizon sample average-feasibility condition, which will prove useful to establish the optimal dynamic ad-allocation policy, i.e., to find a (randomized) policy π, such that

$$\liminf_{T\uparrow+\infty} \frac{1}{T}\sum_{t=1}^{T} y_i^j(t|\pi) \geq \alpha_i^j, \text{ for each } i\in\mathcal{N}, j\in\mathcal{M} \tag{4}$$

Note that a similar periodic-review reformulation of service-level constraints has also been adopted in the literature on resource allocation and inventory pooling (e.g. Zhong et al. 2017, Jiang et al. 2023).

4.2. Necessary and Sufficient Condition for Single-Period Feasibility

To obtain the optimal click-through targets that solve ($2\mathcal{SSP}$), we first characterize the necessary and sufficient condition under which the first-stage click-through target vector $\boldsymbol{\alpha}$ is single-period feasible, i.e., (3) holds. We consider the following formulation with a constant objective function:

$$\begin{aligned} &\max_{\pi\in\Pi_{\mathsf{static}}} \; 0 \\ &\quad \text{s.t. } \mathbb{E}[y_i^j(t|\pi)] \geq \alpha_i^j, \text{ for each } i\in\mathcal{N}, j\in\mathcal{M} \end{aligned} \tag{5}$$

Note that, due to stationarity, (5) is regardless of time t. We now characterize when the stochastic program (5) has a feasible solution. We first reformulate (5) as a linear program (LP). Note that the set of deterministic static policies Π_{d} are all the mappings that take a type-j customer to an offer-set in $\mathfrak{S}^j$, which is finite with cardinality $|\Pi_{\mathsf{d}}| = \prod_{j\in\mathcal{M}} |\mathfrak{S}^j|$. Hence, a randomized static policy π_{static} is defined by a probability measure $\mu(\cdot)$ on the finite set Π_{d}, which is essentially a probability simplex in the space $\mathbb{R}^{|\Pi_{\mathsf{d}}|}$.

Under a deterministic static policy $\pi\in\Pi_{\mathsf{d}}$, if a type-j customer arrives, the platform displays an offer-set $S=\pi(j)$ (due to stationarity, we drop the time index t). Thus, the average per-customer-impression number of click-throughs for ad i by type-j customers is given by $p^j\phi_i^j(\pi(j))$. Therefore, (5) can be reformulated as the following LP, the solution to which we denote as $\mu^*(\cdot)$:

$$\begin{aligned} &\max_{\mu(\cdot)} 0 \\ &\text{s.t. } \sum_{\pi\in\Pi_{\mathsf{d}}} \mu(\pi)p^j\phi_i^j(\pi(j)) \geq \alpha_i^j, \text{ for each } i\in\mathcal{N} \text{ and } j\in\mathcal{M} \\ &\qquad \sum_{\pi\in\Pi_{\mathsf{d}}} \mu(\pi) = 1 \\ &\qquad \mu(\pi)\geq 0 \text{ for all } \pi\in\Pi_{\mathsf{d}} \end{aligned} \tag{6}$$

Taking the dual of the LP (6), we obtain that

$$\begin{aligned}
&\min_{\theta_0,\theta_i^j} \{\theta_0 - \sum_{i\in\mathcal{N},j\in\mathcal{M}} \alpha_i^j \theta_i^j\} \\
&\text{s.t.} \sum_{i\in\mathcal{N},j\in\mathcal{M}} p^j \phi_i^j(\pi(j))\theta_i^j - \theta_0 \le 0, \text{ for all } \pi\in\Pi_{\mathsf{d}} \qquad (7)\\
&\qquad \theta_i^j \ge 0 \text{ for each } i\in\mathcal{N} \text{ and } j\in\mathcal{M}
\end{aligned}$$

Note that, in (7), $\theta_i^j \ge 0$ is the shadow price for the click-through target of ad i from customer segment j, $\sum_{\pi\in\Pi_{\mathsf{d}}} \mu(\pi)p^j\phi_i^j(\pi(j)) \ge \alpha_i^j$, whereas θ_0 is the dual-variable for the normalization constraint $\sum_{\pi\in\Pi_{\mathsf{d}}} \mu(\pi) = 1$. We also define $\boldsymbol{\theta} := (\theta_i^j : i\in\mathcal{N}, j\in\mathcal{M})$.

We note that the objective function of the primal formulation (6) is a constant 0, and there exists a feasible solution $\theta_0 = 0$ and $\theta_i^j = 0$ (for all i and j) to the dual formulation (7) with an objective value equal to 0. By strong duality, (6) has a feasible solution if and only if the optimal objective function value of (7) is nonnegative. By (7), the minimal objective function value of (7) can be obtained at the smallest feasible θ_0, i.e., $\max_{\pi\in\Pi_{\mathsf{d}}} \sum_{i\in\mathcal{N},j\in\mathcal{M}} p^j\phi_i^j(\pi(j))\theta_i^j$ based on the first set of constraints in (7). Combining the aforementioned two observations, (6) is feasible if and only if

$$\min_{\theta_i^j\ge 0} \left\{ \max_{\pi\in\Pi_{\mathsf{d}}} \sum_{i\in\mathcal{N},j\in\mathcal{M}} p^j\phi_i^j(\pi(j))\theta_i^j - \sum_{i\in\mathcal{N},j\in\mathcal{M}} \alpha_i^j\theta_i^j \right\} \ge 0$$

which is equivalent to

$$\max_{\pi\in\Pi_{\mathsf{d}}} \sum_{i\in\mathcal{N},j\in\mathcal{M}} p^j\phi_i^j(\pi(j))\theta_i^j \ge \sum_{i\in\mathcal{N},j\in\mathcal{M}} \alpha_i^j\theta_i^j \text{ for all } \theta_i^j \ge 0,\ i\in\mathcal{N}, j\in\mathcal{M} \qquad (8)$$

The following theorem summarizes the above argument, and it establishes the necessary and sufficient condition for the click-through target vector $\boldsymbol{\alpha}$.

Theorem 1. (Necessary and Sufficient Condition) *A click-through target vector $\boldsymbol{\alpha}$ is single-period feasible, i.e., there exists a static policy π_{static} such that* (3) *holds, if and only if* (8) *holds.*

Indeed, when $\boldsymbol{\alpha}$ satisfies (8), an optimal dual-vector $\boldsymbol{\theta}^*$ that solves (7) helps characterize the set of *deterministic* static policies over which a primal policy $\mu^*(\cdot)$ (feasible to (6)) randomizes. Specifically, strong duality and the complementary slackness condition imply that for a deterministic policy $\pi\in\Pi_{\mathsf{d}}$ to have a positive weight in a feasible primal policy $\mu^*(\cdot)$, i.e., $\mu^*(\pi) > 0$, it must hold that the first constraint of the dual problem (7) is binding for π, i.e.,

$$\pi(j) \in \arg\max_{S\in\mathfrak{S}^j} \sum_{i\in S} p^j\theta_i^j\phi_i^j(S) = \arg\max_{S\in\mathfrak{S}^j} p^j \sum_{i\in S} \theta_i^j\phi_i^j(S) = \arg\max_{S\in\mathfrak{S}^j} \sum_{i\in S} \theta_i^j\phi_i^j(S)$$

Note that the left-hand side of inequality (8) can be viewed as a personalized offer-set optimization problem. Specifically, for each customer type j, we seek to provide an offer-set that maximizes the total revenue from this customer type with his or her per-click revenue of ad i set to θ_i^j, i.e.,

$$S^*(\boldsymbol{\theta}|j) = \arg\max_{S \in \mathfrak{S}^j} \sum_{i \in S} \theta_i^j \phi_i^j(S) \tag{9}$$

For tied solutions, an arbitrary offer-set that solves (9) with the smallest cardinality is displayed to the type-j customer so that $S^*(\boldsymbol{\theta}|j)$ is uniquely determined for a given $\boldsymbol{\theta}$. Given a vector $\boldsymbol{\theta}$, we denote the deterministic policy generated by (9) as $\pi_{\boldsymbol{\theta}}$ (hence, $\pi_{\boldsymbol{\theta}}(j) = S^*(\boldsymbol{\theta}|j)$).

Given the dual-vector $\boldsymbol{\theta}$, we define $g(\boldsymbol{\theta}) := \sum_{i \in \mathcal{N}, j \in \mathcal{M}} p^j \theta_i^j \phi_i^j(S^*(\boldsymbol{\theta}|j))$, which is the left-hand side of (8). Hence, we obtain an equivalent necessary and sufficient condition for the feasibility of click-through targets $\boldsymbol{\alpha}$:

$$\begin{aligned} &h(\boldsymbol{\alpha}) \geq 0 \\ \text{where } &h(\boldsymbol{\alpha}) := \min_{\boldsymbol{\theta} \geq \mathbf{0}} \left\{ g(\boldsymbol{\theta}) - \sum_{i \in \mathcal{N}, j \in \mathcal{M}} \alpha_i^j \theta_i^j : \theta_i^j \geq 0, \text{ for each } i \in \mathcal{N}, j \in \mathcal{M} \right\} \end{aligned} \tag{10}$$

Because $g(\boldsymbol{\theta})$ is the maximum of a family of linear functions, it is jointly convex in $\boldsymbol{\theta}$. Therefore, checking the feasibility of the two-stage stochastic program (2$\mathcal{SSP}$) is reduced to minimizing a convex function $g(\boldsymbol{\theta}) - \sum_{i \in \mathcal{N}, j \in \mathcal{M}} \alpha_i^j \theta_i^j$ over the quadrant $\{\theta_i^j \geq 0 : i \in \mathcal{N}, j \in \mathcal{M}\}$. Hence, as long as the personalized offer-set optimization problem (9) is tractable (i.e., the customer click-throughs follow independent, MNL, nested-MNL, or generalized attraction choice models), one could numerically check the feasibility of the click-through targets $\boldsymbol{\alpha}$. By (10), $h(\boldsymbol{\alpha})$ is the minimum of a family of linear functions (in $\boldsymbol{\alpha}$), so it is jointly concave in $\boldsymbol{\alpha}$.

With the characterization of the necessary and sufficient condition (10) for the feasibility of click-through targets $\boldsymbol{\alpha}$ in the second stage, we are now ready to reformulate the two-stage stochastic program (2$\mathcal{SSP}$) as the following single-stage (*deterministic*) convex program to obtain the optimal target vector:

$$\begin{aligned} \max_{\boldsymbol{\alpha} \geq \mathbf{0}} \ & \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) \\ \text{s.t. } & h(\boldsymbol{\alpha}) \geq 0, \\ & b_i \sum_{j \in \mathcal{M}} \alpha_i^j \leq \frac{B_i}{T}, \text{ for each } i \in \mathcal{N}, \\ & \sum_{j \in \mathcal{C}} \alpha_i^j \geq \frac{\eta_i^{\mathcal{C}}}{T}, \text{ for each } i \in \mathcal{N} \text{ and } \mathcal{C} \in \mathfrak{K}_i \end{aligned} \tag{$\mathcal{OTP}$}$$

Of particular importance is a special case of ($\mathcal{OTP}$) where customer click-throughs follow the MNL choice model, i.e., for any $i \in S$,

$$\phi_i^j(S) = \frac{v_i^j}{1 + \sum_{i' \in S} v_{i'}^j} \tag{11}$$

where $v_i^j > 0$ is the attractiveness of ad i to type-j customers. We demonstrate in the following proposition that if customer click-throughs follow the MNL choice model (11) with the cardinality constraint for any offer-set displayed to a customer (i.e., $|S(t)| \le K$ for some K), the optimal target problem $(\mathcal{OTP})$ can be simplified to a compact convex program with a few linear constraints.[7]

Proposition 1. *If customer click-throughs follow the MNL choice model* (11) *and the set of all feasible offer-sets is* $\mathfrak{S}^j = \{S \subset \mathcal{N} : |S| \le K\}$ *for each* $j \in \mathcal{M}$*, the first-stage convex program* $(\mathcal{OTP})$ *can be simplified to the following one:*

$$\begin{aligned} \max_{\boldsymbol{\alpha} \ge \mathbf{0}} \ & \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) \\ s.t. \ & \sum_{i' \in \mathcal{N}} \alpha_{i'}^j + \frac{\alpha_i^j}{v_i^j} \le p^j, \ \text{for each } i \in \mathcal{N}, j \in \mathcal{M}, \\ & \sum_{i \in \mathcal{N}} \alpha_i^j + \frac{1}{K} \sum_{i \in \mathcal{N}} \frac{\alpha_i^j}{v_i^j} \le p^j, \ \text{for each } j \in \mathcal{M}, \\ & b_i \sum_{j \in \mathcal{M}} \alpha_i^j \le \frac{B_i}{T}, \ \text{for each } i \in \mathcal{N}, \\ & \sum_{j \in \mathcal{C}} \alpha_i^j \ge \frac{\eta_i^{\mathcal{C}}}{T}, \ \text{for each } i \in \mathcal{N} \text{ and } \mathcal{C} \in \mathfrak{K}_i \end{aligned} \qquad (\mathcal{OTP} - \mathcal{MNL})$$

Proposition 1 shows that the number of linear constraints for the convex program $(\mathcal{OTP} - \mathcal{MNL})$ is $\mathcal{O}(mn)$ (instead of exponential in m and n), which ensures its tractability.

Definition 1. We say that a click-through target vector $\boldsymbol{\alpha} \in \mathbb{R}_+^{nm}$ is *feasible* if it is a feasible solution to $(\mathcal{OTP})$.

By definition, if $\boldsymbol{\alpha}$ is *feasible*, then it is *single-period feasible*. Throughout this paper, we assume the *feasible* region is nonempty, so an optimal solution to $(\mathcal{OTP})$ exists, which we denote as $\boldsymbol{\alpha}^*$. Also, we denote $\mathcal{V}_{\mathsf{CT}}^* := \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}^*)$ as the optimal objective function value of $(\mathcal{OTP})$. Thus, $\boldsymbol{\alpha}^*$ is the "optimal" click-through target vector for our reformulated ad-allocation problem. According to Theorem 1, $h(\boldsymbol{\alpha})$ defined by (10) being nonnegative provides a necessary and sufficient condition for the click-through targets, $\boldsymbol{\alpha}$, to be obtainable in the expected sense $\mathcal{V}_{\mathsf{CT}}^*$, which proves to be an upper bound of the optimal FV for the original problem, $\mathcal{V}^*$ (see Theorem 3 below). The convex program formulation $(\mathcal{OTP})$, therefore, characterizes the optimal click-through target vector $\boldsymbol{\alpha}^*$ and the associated optimal (relaxed) per-customer-impression FV in the expected sense. However, the following two critical questions remain to be addressed.

- **Achieving $\boldsymbol{\alpha}^*$:** How should we display the offer-sets upon the arrival of each customer to achieve the optimal click-through targets $\boldsymbol{\alpha}^*$?

[7] In Appendix E, we refer to additional insights on the feasible region of the click-through targets. In Appendix F, we show that if customer click-throughs follow the independent choice model (which is widely adopted in practice; see, e.g., Feldman et al. 2022) or the generalized attraction choice model (which is more general than MNL; see, e.g., Luce 2012, Gallego et al. 2015), the optimal target problem $(\mathcal{OTP})$ can also be simplified to tractable convex programs.

- **Optimality of achieving $\boldsymbol{\alpha}^*$:** Will the offer-set display strategy achieving $\boldsymbol{\alpha}^*$ suffice to obtain the true (unrelaxed) optimal value of the original problem $(\mathcal{OP})$, i.e., $\mathcal{V}^*$?

The rest of this paper is devoted to addressing both questions.

5. Algorithms for Advertisement Allocation Optimization

In this section, we develop an offer-set allocation algorithm under the TTD framework to address the ad-allocation problem $(\mathcal{OP})$ based on the solution to the click-through target model, $\boldsymbol{\alpha}^*$. Specifically, we propose an adaptive offer-set policy that meets the optimal click-through targets $\boldsymbol{\alpha}^*$, and we demonstrate that our proposed algorithm is asymptotically optimal as the problem size scales to infinity. If only a compromised solution can be obtained for the optimal target problem $(\mathcal{OTP})$, the algorithm will achieve the same (asymptotic) optimality gap as that in the optimal target problem, suggesting the robustness of our approach. Our proposed algorithm has better scalability than Fluid-based approaches, and it can deplete the budget of each ad more smoothly throughout the horizon, which is highly desirable for the advertising business in practice.

5.1. Debt-Weighted Advertisement Allocation Policy

By our two-stage stochastic program (re)formulation of the ad-allocation problem, $(2\mathcal{SSP})$, once we solve the optimal click-through target vector $\boldsymbol{\alpha}^*$, the problem is reduced to devising a randomized offer-set algorithm to achieve $\boldsymbol{\alpha}^*$. To this end, one may solve the primal-dual problems (5) and (7) with $\boldsymbol{\alpha} = \boldsymbol{\alpha}^*$ to obtain a feasible randomized policy that achieves $\boldsymbol{\alpha}^*$. This approach, though intuitive, may be computationally prohibitive, because the primal LP (6) has $\mathcal{O}(m2^n)$ decision variables and $\mathcal{O}(mn)$ constraints. Therefore, we resort to a data-driven algorithm to generate the random dual-vector $\boldsymbol{\theta}(t)$ upon the arrival of each customer t, based on which we adaptively customize the appropriate ad offer-set $S^*(\boldsymbol{\theta}(t)|j(t))$. Algorithm 1 below presents our policy. We refer to the DWO policy (Algorithm 1) initialized with the click-through target vector $\boldsymbol{\alpha}$ as the DWO-$\boldsymbol{\alpha}$ policy, denoted by $\pi_{\mathsf{DWO}}(\boldsymbol{\alpha})$. Of particular importance is the DWO-$\boldsymbol{\alpha}^*$ policy, where the platform solves the optimal target problem $(\mathcal{OTP})$ *offline* to obtain the optimal click-through target vector $\boldsymbol{\alpha}^*$, and then implements $\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)$ *online* to adaptively display a personalized offer-set to each customer. The main result of this section is that the DWO-$\boldsymbol{\alpha}^*$ policy is asymptotically optimal for the original ad-allocation problem $(\mathcal{OP})$.

Algorithm 1 DEBT-WEIGHTED OFFER-SET POLICY $\pi_{\mathsf{DWO}}(\boldsymbol{\alpha})$

Initialize: The click-through target vector $\boldsymbol{\alpha}$ and the initial debts $d_i^j(1) \leftarrow 0$ for all $i \in \mathcal{N}$ and $j \in \mathcal{M}$.

For each customer $t = 1, 2, \cdots T$**:**

1: Observe the customer type $j(t)$.

2: Display the offer-set

$$S^*(\boldsymbol{d}(t)|j(t)) := \arg\max_{S \in \mathfrak{S}^j} \sum_{i \in S} (d_i^{j(t)}(t))^+ \phi_i^{j(t)}(S) \tag{12}$$

to customer t, where $\boldsymbol{d}(t) = (d_i^j(t) : i \in \mathcal{N}, j \in \mathcal{M})$ is the realized debt vector upon the arrival of customer t. For tied solutions, an arbitrary offer-set that solves (12) with the smallest cardinality is displayed to the type-j customer.

3: Observe the customer click-throughs $(y_i^j(t) : i \in \mathcal{N}, j \in \mathcal{M})$ and collect the advertising value $\sum_{i \in \mathcal{N}} r_i^{j(t)} y_i^{j(t)}$. Remove any offer-set containing ad i with $\left(\sum_j \sum_{\tau \le t} y_i^j(\tau)\right) b_i \ge B_i$ (i.e., the budget of ad i has already been depleted) from $\mathfrak{S}^j$ for all j hereafter.

4: Update the debt $d_i^j(t+1) \leftarrow d_i^j(t) + \alpha_i^j - y_i^j(t)$ for all $i \in \mathcal{N}$ and $j \in \mathcal{M}$.

A few remarks are in order with respect to Algorithm 1. First, the DWO policy displays the offer-set to each customer based on the offer-set optimization problem (9). This standard personalized offer-set optimization problem is tractable so the offer-set $S^*(\boldsymbol{d}(t)|j(t))$ can be efficiently obtained for a broad class of choice models: independent, MNL, nested MNL, and generalized attraction. For tied solutions, an arbitrary offer-set that solves (12) with the smallest cardinality is displayed to the type-j customer so that $S^*(\boldsymbol{d}(t)|j(t))$ is uniquely defined. Hence, any ad i with $d_i^j(t) \le 0$ will not be offered to customer t with type j.

Second, we call Algorithm 1 the DWO policy, because the offer-set optimization is weighted by the "debt" of each customer-advertisement pair for customers $\{1, 2, ..., t-1\}$. Note that $(t-1)\alpha_i^j$ is the total click-through target of ad i by type-j customers until the start of time t, whereas $\sum_{\tau=1}^{t-1} y_i^j(\tau)$ is the total realized click-throughs by then. Therefore, $\left(d_i^j(t)\right)^+ = \max\left((t-1)\alpha_i^j - \sum_{\tau=1}^{t-1} y_i^j(\tau), 0\right)$ is the total "debt" owed by the platform to the click-through target associated with ad i and customer type j when deciding the offer-set displayed to customer t. For a feasible click-through target vector $\boldsymbol{\alpha}$, we can also view the debt process $\{\boldsymbol{d}(t) : t \ge 1\}$ as a data-driven adaptive way to generate the random dual-vector $\boldsymbol{\theta}$, which prescribes a feasible randomized policy $\pi = \pi_{\boldsymbol{\theta}}$.

Third, note that the debts at the start of time t, $\boldsymbol{d}(t)$ only depend on $\mathcal{H}_{t-1}$ and are independent of any information revealed on or after time t. Finally, under the MNL, independent, and generalized attraction choice models, the offer-set in Step 2 can be obtained efficiently in a real-time manner. For example, Feldman et al. (2022) demonstrate that, for the MNL choice model, the single-period assortment optimization problem can be solved with the running time $\mathcal{O}(n^2)$ (where n is the number of products) and have successfully deployed their algorithm on one of Alibaba's large-scale online retailing platforms (Tmall).

5.2. Asymptotic Analysis

In this subsection, we will establish that the $\mathsf{DWO}\text{-}\boldsymbol{\alpha}^*$ policy can achieve the optimal FV for the original ad-allocation problem ($\mathcal{OP}$) asymptotically. Before demonstrating the optimality of the $\mathsf{DWO}\text{-}\boldsymbol{\alpha}^*$ policy, we first introduce the asymptotic regime where the problem size scales up to infinity. Specifically, we denote a family of ad-allocation problems with the budget for each ad i, $B_i(\gamma) := B_i\gamma$, the click-through requirement for ad i and customer-type set $\mathcal{C} \in \mathfrak{K}_i$, $\eta_i^{\mathcal{C}}(\gamma) = \eta_i^{\mathcal{C}}\gamma$, and the planning horizon length $T(\gamma) := T\gamma$, as $\mathcal{OP}(\gamma)$, where $\gamma > 0$ is a scaling parameter of problem size. Hence, the original problem ($\mathcal{OP}$) is equivalent to $\mathcal{OP}(1)$. For the problem $\mathcal{OP}(\gamma)$ and a policy $\pi \in \Pi$, we denote $\mathcal{E}(\pi|\gamma)$ as the expected efficiency, $\mathcal{F}(\pi|\gamma)$ as the expected fairness, and $\mathcal{V}(\pi|\gamma) = \mathcal{E}(\pi|\gamma) + \lambda\mathcal{F}(\pi|\gamma)$ as the expected FV generated by π in $\mathcal{OP}(\gamma)$. Furthermore, $\mathcal{V}^*(\gamma) := \max_{\pi\in\Pi} \mathcal{V}(\pi|\gamma)$ denotes the optimal expected FV for $\mathcal{OP}(\gamma)$. Note that the market-size scaling factor γ does not affect the feasibility of a click-through target vector $\boldsymbol{\alpha}$, nor does it change the two-stage stochastic program reformulation ($2\mathcal{SSP}$) or the target problem reformulation ($\mathcal{OTP}$).

We first establish that, for any *feasible* click-through target vector $\boldsymbol{\alpha}$, the $\mathsf{DWO}\text{-}\boldsymbol{\alpha}$ policy exactly achieves $\boldsymbol{\alpha}$ in $\mathcal{OP}(\gamma)$ as the problem size γ scales to infinity.

THEOREM 2. *If $\boldsymbol{\alpha}$ is feasible, i.e., all constraints of ($\mathcal{OTP}$) are satisfied, then we have:*

$$\lim_{\gamma\uparrow+\infty} \frac{1}{T(\gamma)} \sum_{t=1}^{T(\gamma)} y_i^j(t|\pi_{\mathsf{DWO}}(\boldsymbol{\alpha})) = \alpha_i^j \text{ almost surely for all } i \in \mathcal{N} \text{ and } j \in \mathcal{M} \tag{13}$$

Theorem 2 is the central technical result of this paper—it is an important stepping stone on the way to proving the asymptotic optimality of the $\mathsf{DWO}\text{-}\boldsymbol{\alpha}^*$ policy. Interestingly, as long as this policy is initiated with a *feasible* click-through target vector $\boldsymbol{\alpha}$, it will not only achieve click-through levels *at least as high as* these targets (i.e., (4)) but also *exactly approach* them (i.e., (13)). Adopting a coupling argument, the proof of Theorem 2 (see Appendix D for details) demonstrates that if the problem size γ scales up to infinity, the $\mathsf{DWO}\text{-}\boldsymbol{\alpha}$ policy will *not* exhaust the budget of any ad and, thus, will secure the click-through targets $\boldsymbol{\alpha}$. Therefore, under our proposed DWO policy, the click-through requirements $\{\eta_i^{\mathcal{C}}(\gamma) : i \in \mathcal{N}, \mathcal{C} \in \mathfrak{K}_i\}$ can be satisfied almost surely, instead of in expectation, in the asymptotic regime.

Based on Theorem 2, one may conjecture that, if the click-through target vector $\boldsymbol{\alpha}$ is optimally chosen (i.e., as the solution to the optimal target problem ($\mathcal{OTP}$), $\boldsymbol{\alpha}^*$), the $\mathsf{DWO}\text{-}\boldsymbol{\alpha}^*$ policy could achieve the optimal FV for the original ad-allocation problem, $\mathcal{V}^*$. The main result of this section is that the following theorem validates this conjecture in the asymptotic regime and quantifies the nonasymptotic optimality gap of the policy.

Theorem 3. *The* $\mathsf{DWO}\text{-}\boldsymbol{\alpha}^*$ *policy is asymptotically optimal, i.e.,*

$$\lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma)=\lim_{\gamma\uparrow+\infty}\mathcal{V}^*(\gamma)=\mathcal{V}^*_{\mathsf{CT}} \tag{14}$$

Furthermore, the optimal objective function value of the first-stage click-through target optimization ($\mathcal{OTP}$) *is an upper bound for the original problem* ($\mathcal{OP}$) *in the nonasymptotic regime, i.e., for any* $\gamma>0$,

$$\mathcal{V}^*_{\mathsf{CT}}\geq\mathcal{V}^*(\gamma), \tag{15}$$

and the optimality gap the $\mathsf{DWO}\text{-}\boldsymbol{\alpha}^*$ *policy is of order* $\mathcal{O}(\gamma^{-\frac{1}{2}})$, *i.e., there exists a constant* $\mathcal{C}>0$, *such that for any* $\gamma>0$,

$$\mathcal{V}^*(\gamma)-\mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma)\leq\frac{\mathcal{C}}{\sqrt{\gamma}}. \tag{16}$$

Theorem 3 proves that the $\mathsf{DWO}\text{-}\boldsymbol{\alpha}^*$ policy generated by our TTD framework is asymptotically optimal when the ad budgets, the click-through requirements, and the time-horizon length all scale up to infinity at the same rate. In particular, the optimal expected FV of $\mathcal{OP}(\gamma)$ is identical to the optimal FV of the optimal target problem ($\mathcal{OTP}$) asymptotically, and the former is upper-bounded by the latter in the nonasymptotic regime. Such equivalence suggests that our reformulation in Section 4.1 is an asymptotically equivalent relaxation of the original problem. Moreover, we show that the $\mathsf{DWO}\text{-}\boldsymbol{\alpha}^*$ policy can achieve the optimal expected FV at a convergence rate of $\mathcal{O}(\gamma^{-\frac{1}{2}})$. The proof of Theorem 3 relies on a delicate application of Theorem 2, which shows that the $\mathsf{DWO}\text{-}\boldsymbol{\alpha}^*$ policy generated by our TTD framework achieves the optimal FV of the click-through target optimization (i.e., $\lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma)=\mathcal{V}^*_{\mathsf{CT}}$). Another critical step to establish Theorem 3 is to find an asymptotically optimal static policy for the original problem $\mathcal{OP}(\gamma)$, which can be constructed with the auxiliary Fluid convex program ($\mathcal{OP}_{\mathsf{Fluid}}$), i.e., Proposition 2. We relegate the detailed discussions to Section 5.3.

In the existing literature, debt-weighted algorithms in the contexts of resource pooling, (e.g., Zhong et al. 2017, Jiang et al. 2023), and process flexibility (e.g., Lyu et al. 2019) have been validated to satisfy feasibility conditions of different formats. Jiang et al. (2023) further provide the optimality guarantee of this policy. Our study not only provides the *optimality* guarantee of a *debt-weighted* policy in addition to its *feasibility/approachability*, but also develops the comprehensive TTD framework that solves a complex ad-allocation problem to (asymptotic) optimality, subject to the ad budget constraints and the click-through requirements. Furthermore, to our best knowledge, we are also the first in the literature to study the dynamic assortment/offer-set optimization problem through the lens of a debt-weighted algorithm.

5.3. Discussions

Fluid **Benchmark.** To prove Theorem 3, one needs to construct an asymptotically optimal static policy for the origianl problem $(\mathcal{OP})$. To find such a policy, we consider an auxiliary Fluid convex program $(\mathcal{OP}_{\mathsf{Fluid}})$ and establish its intrinsic connections and, therefore, asymptotic equivalence to the original problem $(\mathcal{OP})$. We also note that the convex program is a generalization of the standard CDLP approach (Liu and Van Ryzin 2008). Specifically, the auxiliary Fluid problem is defined as the convex program $(\mathcal{OP}_{\mathsf{Fluid}})$:

$$\begin{aligned}
\max_{\boldsymbol{z}}\ & \mathcal{V}_{\mathsf{Fluid}}(\boldsymbol{z}) := \sum_{i\in\mathcal{N},j\in\mathcal{M},S\in\mathfrak{S}^j} r_i^j p^j \phi_i^j(S) z^j(S) + \lambda F(\boldsymbol{\zeta}) \\
\text{s.t.}\ & \sum_{j\in\mathcal{M},S\in\mathfrak{S}^j} b_i p^j \phi_i^j(S) z^j(S) \le \frac{B_i}{T} \text{ for each } i\in\mathcal{N} \\
& \sum_{j\in\mathcal{C},S\in\mathfrak{S}^j} p^j \phi_i^j(S) z^j(S) \ge \frac{\eta_i^{\mathcal{C}}}{T} \text{ for each } i\in\mathcal{N} \text{ and } \mathcal{C}\in\mathfrak{K}_i \\
& \sum_{S\in\mathfrak{S}^j} z^j(S) \le 1 \text{ for each } j\in\mathcal{M} \\
& z^j(S)\ge 0 \text{ for each } j\in\mathcal{M}, S\in\mathfrak{S}^j \\
& \boldsymbol{\zeta}\in\mathbb{R}^{nm},\ \text{with } \zeta_i^j = \sum_{S\in\mathfrak{S}^j} p^j z^j(S)\phi_i^j(S)
\end{aligned} \qquad (\mathcal{OP}_{\mathsf{Fluid}})$$

We denote the solution to $(\mathcal{OP}_{\mathsf{Fluid}})$ as $\boldsymbol{z}^*$, and the associated optimal objective function value as $\mathcal{V}^*_{\mathsf{Fluid}} = \mathcal{V}_{\mathsf{Fluid}}(\boldsymbol{z}^*)$. It is self-evident from the formulation of $(\mathcal{OP}_{\mathsf{Fluid}})$ that $z^j(S)$ is the probability of displaying offer-set S to a type-j customer upon his or her arrival, whereas $\zeta_i^j = \sum_{S\in\mathfrak{S}^j} p^j z^j(S)\phi_i^j(S)$ is the expected per-customer-impression click-throughs of ad i by type-j customers. A vector $\boldsymbol{z} = (z^j(S) : j\in\mathcal{M}, S\in\mathfrak{S}^j)$ feasible for $(\mathcal{OP}_{\mathsf{Fluid}})$ naturally induces a randomized, static policy for the original problem $(\mathcal{OP})$, which displays offer-set $S\in\mathfrak{S}^j$ to a customer of type $j\in\mathcal{M}$ with probability $z^j(S)$. Whenever at least one ad runs out of budget, the policy offers nothing to each arriving customer afterwards. We refer to this policy as the Fluid-$\boldsymbol{z}$ policy, denoted by $\pi_{\mathsf{Fluid}}(\boldsymbol{z})$. Note that policy $\pi_{\mathsf{Fluid}}(\boldsymbol{z})$ does not fully utilize the remaining budgets of non-depleted ads. To strengthen the performance of this policy, we will slightly adjust its implementation in our numerical experiments. The detailed discussions of the adjustment are deferred to Section 6.

Indeed, there are intrinsic connections between the optimal target problem $(\mathcal{OTP})$ and the Fluid convex program $(\mathcal{OP}_{\mathsf{Fluid}})$. We can always construct a feasible (resp. optimal) click-through target vector in $(\mathcal{OTP})$ from any feasible (resp. optimal) offer-set assignment probabilities in $(\mathcal{OP}_{\mathsf{Fluid}})$. For any $\boldsymbol{z}$ feasible to $(\mathcal{OP}_{\mathsf{Fluid}})$, we define $\hat{\boldsymbol{\alpha}}(\boldsymbol{z})\in\mathbb{R}^{nm}$, where $\hat{\alpha}_i^j(\boldsymbol{z}) = \sum_{S\in\mathfrak{S}^j} p^j z^j(S)\phi_i^j(S)$.

LEMMA 2. *Assume that $\boldsymbol{z}$ is feasible for $(\mathcal{OP}_{\mathsf{Fluid}})$. We have that $\hat{\boldsymbol{\alpha}}(\boldsymbol{z})$ is first-stage feasible and can be achieved by policy $\pi_{\mathsf{Fluid}}(\boldsymbol{z})$, i.e., $\mathbb{E}[y_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}))] = \hat{\alpha}_i^j(\boldsymbol{z})$. Furthermore, $\mathcal{V}_{\mathsf{CT}}(\hat{\boldsymbol{\alpha}}(\boldsymbol{z})) = \mathcal{V}_{\mathsf{Fluid}}(\boldsymbol{z})$. In particular, $\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)$ is an optimal solution to $(\mathcal{OTP})$ with $\mathbb{E}[y_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*))] = \hat{\alpha}_i^j(\boldsymbol{z}^*)$.*

We are now ready to demonstrate that, as the problem size γ scales to infinity, both the original problem $\mathcal{OP}(\gamma)$ and the Fluid convex program $(\mathcal{OP}_{\mathsf{Fluid}})$ have the same optimal (expected) per-customer-impression FV, which is also identical to the one generated by the Fluid-z^* policy in $\mathcal{OP}(\gamma)$. Formally, the following proposition establishes these equivalences and implies the asymptotic optimality of a static policy $\pi_{\mathsf{Fluid}}(z^*)$.

Proposition 2. *The following inequalities hold:*

$$\lim_{\gamma\uparrow+\infty}\mathcal{V}^*(\gamma)\geq\lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{Fluid}}(z^*)|\gamma)=\mathcal{V}^*_{\mathsf{Fluid}}\geq\lim_{\gamma\uparrow+\infty}\mathcal{V}^*(\gamma) \tag{17}$$

Therefore, all inequalities in (17) *hold as equalities.*

As a stepping stone to prove Theorem 3, Proposition 2 shows that the static Fluid-z^* policy generated by $(\mathcal{OP}_{\mathsf{Fluid}})$ is asymptotically optimal for the original problem $\mathcal{OP}(\gamma)$. Hence, we will also use this policy and its re-solving variants as the benchmarks in our numerical experiments to evaluate our proposed DWO-α^* policy in Section 6. Note that the equivalences in (17) generalize Proposition 1 of Liu and Van Ryzin (2008) to our setting with algorithmic fairness, personalized offer-sets, and click-through requirements.

Computational efficiency. Our proposed DWO-α^* policy generated by our TTD framework involves two steps. The first step solves the optimal target problem $(\mathcal{OTP})$ to obtain the optimal click-through target vector α^* offline, and the second step implements the second-stage DWO display procedure online. We now discuss the computational efficiency of the two steps separately, starting from the second-stage online implementation.

Second-stage online implementation. To implement Algorithm 1 online, given any click-through target vector α, the bottleneck is to solve a single-period offer-set optimization problem (9) upon the arrival of each customer t. Standard results in the assortment-optimization literature suggest that if customers follow a wide range of commonly used choice models—such as the independent choice (Feldman et al. 2022), MNL (Rusmevichientong et al. 2010, Davis et al. 2013), nested MNL (Davis et al. 2014), and generalized attraction (Luce 2012, Gallego et al. 2015) choice models—the personalized offer-set optimization (9) can be solved efficiently. Therefore, the second-stage online implementation of the DWO-α is computationally efficient as long as the single-period offer-set optimization is tractable, which is generally the case for choice models commonly used in practice.

First-stage convex program. Section 4.2 and Appendix F show that if the customer choices follow the MNL, independent, and generalized attraction choice models, the optimal target problem $(\mathcal{OTP})$ can be greatly simplified to a convex program with a few linear constraints, which can be solved efficiently in general. We emphasize that the MNL, independent, and generalized attraction choice models are all widely used in practice. If customers follow a general choice model, Lemma 2 implies that $(\mathcal{OTP})$ shares the same computational complexity as the Fluid convex program $(\mathcal{OP}_{\mathsf{Fluid}})$. To see this, note that for any solution to $(\mathcal{OP}_{\mathsf{Fluid}})$, $\boldsymbol{z}^*$, we can construct a click-through target vector $\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)$ that is feasible and optimal for $(\mathcal{OTP})$. Therefore, as long as the auxiliary Fluid convex program $(\mathcal{OP}_{\mathsf{Fluid}})$ is computationally tractable, we can efficiently obtain an optimal solution to $(\mathcal{OTP})$ as well. In general, the DWO-$\boldsymbol{\alpha}^*$ policy solves $(\mathcal{OTP})$ offline only once at the beginning of planning horizon, which is generally tractable in most applications. Indeed, Table 1 in Section 6 shows that our DWO-$\boldsymbol{\alpha}^*$ policy is much more scalable than the Fluid-based benchmarks commonly adopted in the literature.

If the auxiliary Fluid convex program $(\mathcal{OP}_{\mathsf{Fluid}})$ is intractable, obtaining an *optimal* click-through target vector $\boldsymbol{\alpha}^*$ may be prohibitive. However, we can still identify a *feasible* click-through target vector $\boldsymbol{\alpha}$ by applying Theorem 1. Once a feasible $\boldsymbol{\alpha}$ is found, as long as the single-period offer-set optimization (9) is tractable, the second-stage online implementation of the DWO-$\boldsymbol{\alpha}$ policy should be tractable as well. Furthermore, the DWO-$\boldsymbol{\alpha}$ policy achieves the same asymptotic FV as $\mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha})$, as shown in the following proposition.

Proposition 3. *If $\boldsymbol{\alpha}$ is first-stage feasible, i.e., it is a feasible solution to $(\mathcal{OTP})$, then we have:*

$$\lim_{\gamma\uparrow+\infty} \mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha})|\gamma) = \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) \tag{18}$$

The key implication from Proposition 3 is that, in the *asymptotic regime* where the problem size γ scales to infinity, the second-stage online implementation of the DWO-$\boldsymbol{\alpha}$ policy will not incur any additional optimality loss on top of that from a feasible suboptimal click-through target vector $\boldsymbol{\alpha}$ for the first-stage target optimization, i.e., $\lim_{\gamma\uparrow+\infty}\mathcal{V}^*(\gamma) - \lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha})|\gamma) = \mathcal{V}^*_{\mathsf{CT}} - \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha})$.

Comparison with existing algorithms. It is useful to compare the DWO-$\boldsymbol{\alpha}^*$ policy with relevant algorithms in the existing literature.

Debt-weighted resource-allocation algorithms. As discussed above, the objective of existing debt-weighted algorithms (e.g, Zhong et al. 2017, Lyu et al. 2019, Jiang et al. 2023) is to allocate a centralized resource to satisfy some *feasible* and *exogenous* service-level constraints. The goal of the DWO-$\boldsymbol{\alpha}^*$ policy, however, is to maximize the FV of an online advertising system so that the click-through target vector, which is the counterpart of the service-level constraints in our setting, is *endogenized* in the first-stage of the algorithm. Because we have such a different objective for our policy, we develop a two-stage reformulation to implement the algorithm and take a different path for its analysis, which relies on establishing the (asymptotic) equivalence of different formulations of the problem. Another key difference between our DWO-$\boldsymbol{\alpha}^*$ policy and other debt-weighted resource-allocation algorithms is that whereas those algorithms can freely control the allocation and consumption of the resources, our policy has to handle the additional complexity of customers' stochastic choice behaviors, which introduces another layer of challenge to controlling the debt process. Moreover, we show that our DWO-$\boldsymbol{\alpha}^*$ policy can work for general resource allocation and other ad-allocation problems, such as Ad Display and AdWords (see, e.g., Mehta et al. 2013) in Appendix G.

Fluid *convex program heuristics.* In our setting with a nonlinear fairness term, the standard LP-based heuristics with or without re-solving (e.g., Liu and Van Ryzin 2008, Jasin and Kumar 2012, Bumpensanti and Wang 2020, Balseiro et al. 2023) applied to linear rewards can be extended to similar heuristics based on Fluid convex programs (e.g., ($\mathcal{OP}_{\mathsf{Fluid}}$)). A core advantage of the DWO-$\boldsymbol{\alpha}^*$ policy over the family of Fluid heuristics is that, by the nature of the algorithm to assign a higher weight to an ad-customer pair with a larger debt, the click-through and, thus, the reward process will follow a mean-reverting pattern. Therefore, compared with the Fluid or Fluid-resolving heuristics, our DWO-$\boldsymbol{\alpha}^*$ algorithm can deplete the budget of each ad more smoothly throughout the horizon, which is highly desirable for the advertising business in practice.

For example, Google recommends a "standard" ad-delivery scheme for its advertisers, especially those with a low budget, to avoid exhausting their budgets early.[8] Under this standard delivery scheme, each advertisement can reach customers evenly throughout the day. Furthermore, with the cardinality constraint on the displayed offer-sets, one difficulty of using Fluid or its resolving variants is that the number of variables (i.e., the probability of each offer-set for all customer types) quickly explodes as the number of products increases, even when the choice model is restricted to MNL. The DWO-$\boldsymbol{\alpha}^*$ policy has a better scalability than those Fluid-based heuristics. These advantages of our policy are also confirmed by our numerical comparisons in Section 6.

Inventory-balancing policy. Inventory balancing is another family of algorithms to address the personalized-assortment optimization problem with inventory constraints (see, e.g., Golrezaei et al. 2014). This policy uses the remaining inventory to re-weight the value of each product. The inventory-balancing policy is difficult, if not impossible, to adapt to our setting because, on one hand, it is challenging for this policy to handle the click-through requirements $\{\eta_i^{\mathcal{C}} : i \in \mathcal{N}, \mathcal{C} \in \mathfrak{K}_i\}$ and, on the other hand, it is hard to incorporate the nonlinear fairness metric into the inventory-balanced offer-set optimization problem upon the arrival of each customer. Our DWO-$\boldsymbol{\alpha}^*$ policy circumvents these two challenges under our two-stage framework within which the personalized offer-set optimization problem is reduced to a standard, single-period problem with a reward linear in the number of click-throughs (9). For completeness, in Appendix K, we numerically compare our DWO-$\boldsymbol{\alpha}^*$ policy with the inventory-balancing benchmark in the setting without click-through requirements (i.e., $\mathfrak{K}_i = \emptyset$ for all i) and fairness concerns (i.e., $\lambda = 0$). The numerical results demonstrate that our policy outperforms the inventory-balancing benchmark for all the problem instances examined when the demand-to-supply ratio is not too large. When the demand-to-supply ratio is large, our policy performs fairly well, achieving an average of more than 99% of the theoretical upper-bound in all problem instances.

[8] See https://support.google.com/google-ads/answer/2404248?hl=en.

6. Numerical Experiments

In this section, we numerically evaluate our DWO-$\boldsymbol{\alpha}^*$ policy (simply DWO in this section) generated by our TTD framework for ad-allocation optimization, benchmarked against four Fluid-based heuristics. The first benchmark is the static policy induced by the optimal solution $\boldsymbol{z}^*$ to $(\mathcal{OP}_{\mathsf{Fluid}})$ (see, also, Liu and Van Ryzin 2008), denoted as the *fluid-approximation policy* or the Fluid-$\boldsymbol{z}^*$ policy.[9] In each period, we randomly display offer-set S to a customer of type j with probability $z^{j*}(S)$, which is a solution to the Fluid convex program $(\mathcal{OP}_{\mathsf{Fluid}})$. Once the budget of an ad is depleted, it is automatically deleted in any offer-set generated by the Fluid policy. In this case, all the remaining ads in the offer-set will continue to be displayed. This adjustment of the Fluid policy is to enhance budget utilization and, consequently, the performance of the policy. The second benchmark is a re-solving variant of the Fluid policy, denoted as the Fluid *re-solving policy* or the Fluid-R policy, which re-solves the Fluid convex program at evenly spaced time epochs based on the remaining budgets and click-through requirements (see, e.g., Jasin and Kumar 2012). The third benchmark is a refined version of the Fluid-R policy, denoted as the Fluid *infrequent re-solving policy* or the Fluid-I-R policy, under which the resolving time epochs are more carefully designed and are infrequent/sparse at the beginning of the ad campaign (see, e.g., Bumpensanti and Wang 2020). Finally, the fourth benchmark re-solves the Fluid convex program at every period (see, e.g., Balseiro et al. 2023), denoted as the Fluid *every-period re-solving policy* or the Fluid-E-R policy. See Appendix H for the implementation details of the Fluid-R, Fluid-I-R, and Fluid-E-R policies.

The key takeaways from our numerical experiments are summarized as follows: (a) Thanks to the mean-reverting nature of our algorithm, the DWO policy achieves better-than-expected performance in FV, and it delivers much more stable FVs (i.e., with lower standard deviations) than the Fluid-based benchmarks. This is because the debt process of the DWO policy directly steers the click-throughs of each ad from each customer type toward the respective optimal target throughout the ad campaign. (b) The click-through requirements are substantially more likely to be satisfied by our DWO policy than the Fluid-based heuristics (except for the Fluid-E-R policy). This further justifies the efficacy of our algorithm in the presence of click-through requirements. (c) Our modeling TTD framework together with the DWO policy could help achieve promising algorithmic fairness without significantly compromising efficiency. (d) With lower time and space complexities, our DWO policy is more scalable and efficient than the Fluid-based benchmarks. Therefore, our approach is well-suited for tackling the aforementioned challenge to address such algorithmic discrimination/bias in ad-delivery optimization.

We consider an ad-allocation problem with $T = 1{,}000$ customers of five types and 50 ads. The customer-type distribution $(p^1, p^2, \ldots, p^5)$, where $\mathbb{P}[j(t) = j] = p^j$ and $\sum_{j=1}^{5} p^j = 1$, is generated from a

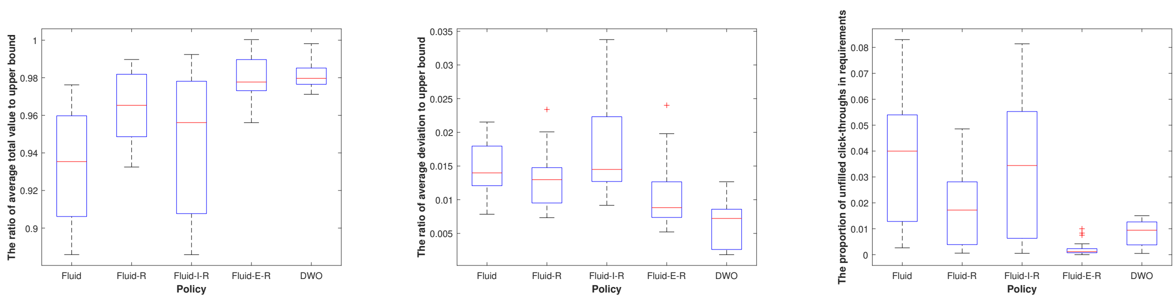
[9] When there is no confusion in the context, we drop $\boldsymbol{z}^*$ and abbreviate it as the Fluid policy.

five-dimensional Dirichlet distribution (see Appendix I for details). We sample the per-click value r_i^j for each ad i by type j independently from a uniform distribution on the interval $[10, 50]$. We model the click-through behavior of the customers using MNL, i.e., for each $i \in S$ and $j \in \mathcal{M}$, $\phi_i^j(S)$ is given by (11). Each ad/customer-type pair is associated with an attraction index v_i^j. For ad i and customer type j, let $v_i^j := \exp(u_i^j)$, where u_i^j is independently sampled from the uniform distribution on the interval $[0, 5]$. We set the cardinality constraint such that the maximum size of an offer-set is 2, i.e., $|S(t)| \leq 2$ for each customer t. The fairness metric we use in the numerical studies is the GMD fairness (24) (in Appendix C) with $\lambda = 10$.

Our first set of numerical studies is based on problem instances generated by systematically varying two focal parameters: (a) the concentration parameter (CP) associated with the proportion of each customer type, and (b) the loading factor (LF), defined as the ratio of total expected demand to total supply. Specifically, the concentration parameters are determined by the parameters of the Dirichlet distribution that we use to generate $(p^1, p^2, \cdots, p^5)$ (see Appendix I for detailed explanations of the CP parameter). The loading factor is the ratio of total user traffic to total affordable traffic with the budgets of the advertisers, namely $LF = T / \sum_{i \in \mathcal{N}} \frac{B_i}{b_i}$. As is clear from their definitions, CP measures the uniformity of the customer-type distribution, while LF measures the tightness of the ad budget. The higher the CP, the more uniform the distribution of customer types; the higher the LF, the tighter the budget constraints for the ad campaigns.

In our experiments, we vary CP in the set $\{0.1, 1, 10, 100\}$ and LF in the set $\{0.5, 0.75, 1, 1.25, 1.5\}$. For each problem instance, we solve the problem $(\mathcal{OTP} - \mathcal{MNL})$ with different per-click values sampled from the same distribution as r_i^j's and without click-through requirements to obtain the solution-optimal targets $\boldsymbol{\alpha}^*$. A click-through requirement $\eta_i^{\{j\}}$, where $\{j\}$ is a singleton set of customer type j, is generated by the product of α_i^{j*} and a random number independently sampled from the $[0, 1]$ uniform distribution for each ad i and each customer type j. We generate 30 sample paths for each problem instance to evaluate the following performance metrics of interest: (1) the ratio between the expected FV and its theoretical upper bound characterized by the solution to the first-stage convex program, $\mathcal{V}^*_{\mathsf{CT}}$; (2) the ratio between the standard deviation of FV and $\mathcal{V}^*_{\mathsf{CT}}$; and (3) the average proportion of unfilled click-through requirements compared with $\eta_i^{\{j\}}$. We use the relative ratios (instead of the absolute values) to make the comparisons clear.

We report the numerical findings as box plots with respect to all problem instances in Figure 2, which clearly illustrates the advantages of our algorithm over the benchmarks in various dimensions. Figure 2(a) demonstrates that our DWO algorithm consistently outperforms all the Fluid-based benchmarks by delivering higher values in the total objective. Figure 2(b) shows that the variability of FV is much lower under our policy than it is under the benchmarks. Finally, Figure 2(c) shows that the DWO algorithm significantly reduces the proportion of unfilled click-through requirements compared with Fluid, Fluid-R, and Fluid-I-R policies, but has more unfilled click-throughs than Fluid-E-R. In short, our proposed DWO algorithm not only generates higher FV than the Fluid-based benchmarks do but also reduces the variability of FV. Note that the re-solving policies can achieve an optimality gap of order $\mathcal{O}(\gamma^{-1})$ under certain regularity conditions (see, e.g., Jasin and Kumar 2012, Bumpensanti and Wang 2020, Balseiro et al. 2023), whereas our DWO policy has a larger provable optimality gap of order $\mathcal{O}(\gamma^{-\frac{1}{2}})$. However, our DWO algorithm outperforms all the Fluid-based benchmarks in the numerical experiments. We propose providing more in-depth analyses of the DWO policy in future research.

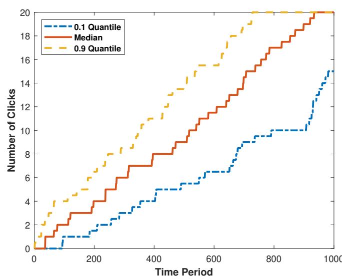


(a) Total Fairness-Adjusted Value (b) Standard Deviation of FV (c) Unfilled Click-Throughs

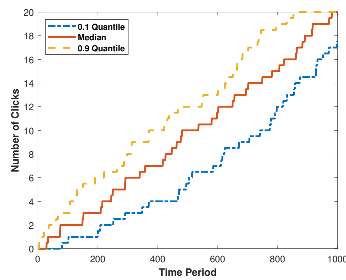
Figure 2 Comparison Between DWO-Based and Fluid-Based Policies

To understand why our DWO algorithm enjoys the great performance illustrated in Figure 2, we also plot the 0.1-, 0.5- (i.e., median) and 0.9- quantiles of the click-through sample-paths of one specific ad for the five approaches we studied (under the problem instance $LF = 1$ and $CP = 100$) in Figure 3. Our numerical experiments make clear that, although all five policies deplete the ad's entire budget for more than 50% of sample paths, the variability of the click-through sample paths (equivalently, the budget-depleting process) through the entire time horizons under the Fluid, Fluid-R, Fluid-I-R, and Fluid-E-R algorithms are much higher than our DWO policy. Furthermore, the Fluid-based approaches all run out of budget long before the end of the ad campaign, while our DWO policy exhausts the budget only toward the very end.

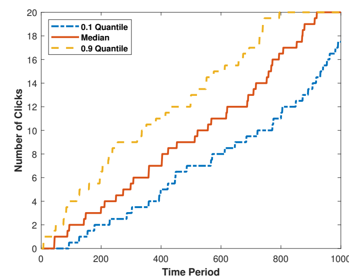
We highlight that such smooth budget depletion of our proposed algorithm should be credited to their mean-reverting pattern driven by the fact that the offer-set displayed in each period is prescribed in accordance with the "debts" owed by the algorithm to the optimal click-through targets. In particular, the ads farther from (resp. closer to) their optimal targets will receive higher (resp. lower) weights when the algorithm is deciding which assortment to display upon the arrival of each user. Thus, such intertemporal pooling leads to the mean-reverting phenomenon of our proposed approach. Note that the Fluid-based policies also exhibit certain mean-reverting property weaker than our DWO policy.[10] Smooth budget depletion is, in practice, a highly desirable property for advertisers that use online advertising platforms—Facebook has even built some API tools that help its clients pace their ad delivery and smooth their budget depletion.[11] Therefore, from a practical perspective, our DWO algorithm may, appealingly, help advertisers and advertising platforms to achieve smoother budget depletion.



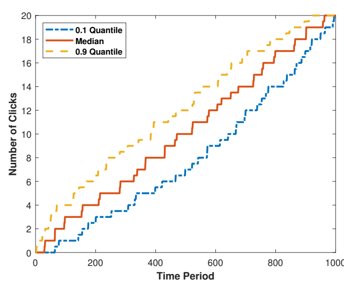
(a) Fluid



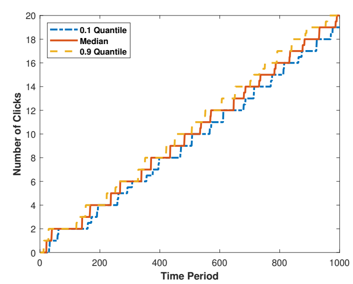
(b) Fluid-R



(c) Fluid-I-R



(d) Fluid-E-R



(e) DWO

Figure 3 The 0.1 Quantiles, Medians, and 0.9 Quantiles in 30 Sample Paths Over Time of Click-Numbers of the Highest-Reward Advertisement With $LF = 1$, and $CP = 100$

In addition to obtaining better performance in most of the cases we examine and much smoother depletion of ad budgets, our DWO policy is more scalable and efficient in both time and space complexities. We carried out our numerical studies by varying the offer-set size constraint K from 2 to 5, and generating 30 samples randomly for each K (under the problem instance $LF = 1$ and $CP = 100$) with other model primitives identical to those of the experiments in Section 6. We conducted the experiment by using Gurobi 10.0.0 within MATLAB R2022b on a 2.10 GHz Intel Core i7-1260P CPU with 32 GB of RAM. Table 1 shows that the average computation time of finding optimal click-through targets (i.e., solving the convex program ($\mathcal{OTP}$)) is approximately 0.01s regardless of the

[10] In Appendix J, we regress the click-through on the per-period debt for all five algorithms. A high per-period debt of an ad-customer pair has a much stronger impact on the potential click-throughs under our DWO policy compared to the Fluid, Fluid-R, Fluid-I-R, and Fluid-E-R algorithms.

[11] See `https://developers.facebook.com/docs/marketing-api/bidding/overview/pacing-and-scheduling`.

value of K, but solving the Fluid convex program ($\mathcal{OP}_{\text{Fluid}}$) is much more time-consuming (Fluid-R, Fluid-I-R and Fluid-E-R are, of course, even slower). In addition, increasing K means exponentially more possible offer-sets for the Fluid, Fluid-R, Fluid-I-R and Fluid-E-R policies, so much more memory and computational time are needed in this case. Table 1 shows that the case of $K = 5$ may even incur an "out of memory" error for the Fluid benchmark. In short, our algorithms enjoy higher scalability than the Fluid-based benchmarks.

Policy	$K=1$	$K=2$	$K=3$	$K=4$	$K=5$
DWO	0.0054s	0.0067s	0.0095s	0.0099s	0.0064s
Fluid	0.0080s	0.0918s	2.5454s	60.1610s	out of memory

Table 1 Comparison of Average Solving Times

Next, we demonstrate the efficiency-fairness trade-off by varying the parameter λ in our setting. Specifically, we consider the problem instance with $\lambda \in \{10^{i_\lambda} : i_\lambda = -1 + 0.1 \times (i-1), i = 1, 2, ..., 31\} \cup \{0\}$ (under the problem instance $LF = 1$ and $CP = 100$). We plot the relationship between the efficiency and Gini fairness in Figure 4 for different values of λ, where the x-axis (resp. y-axis) is the ratio between the expected efficiency (resp. expected GMD fairness) with respect to λ and that with respect to $\lambda = 0$ (i.e., the system is purely efficiency-driven). Our numerical results reveal the trade-off between efficiency and fairness. Importantly, we find that introducing the fairness term in the objective function could substantially reduce the algorithmic bias without much compromising the advertising efficiency. For example, a 1% (resp. 5%) optimality gap in efficiency could reduce about 50% (resp. 90%) of the algorithmic bias.

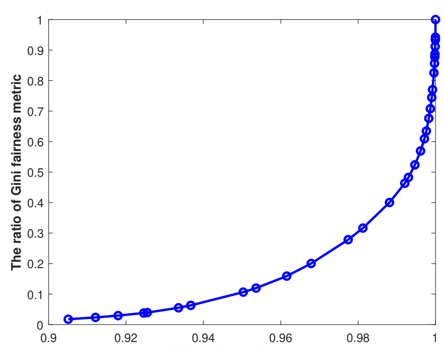


Figure 4 The Trade-Off Between Optimal Efficiency and Gini Fairness

7. Conclusion

The allocation of customer traffic to different ads is a crucial operations decision for online e-commerce platforms to optimize their advertising business. The emerging advocacy for algorithmic fairness of online ad delivery has posed additional challenges for the design of ad-allocation policies. In this paper, we propose a TTD framework comprising a general model and an associated efficient algorithm to study optimal ad allocation under customer choices and algorithmic fairness. Although the original online ad-allocation problem is intractable, we develop an asymptotically equivalent two-stage stochastic program as a surrogate. Furthermore, we propose a simple but effective algorithm—the DWO-$\boldsymbol{\alpha}^*$ policy—which is provably optimal for achieving the maximum FV from advertising in the asymptotic regime. Furthermore, the proposed algorithm gives rise to the mean-reverting pattern of the budget-consumption process and, therefore, achieves smoother budget depletion, which is highly desirable from a practical perspective. Our algorithm also helps substantially improve the fairness of ad allocation for a platform without compromising its efficiency much.

References

Ali, Muhammad, Piotr Sapiezynski, Miranda Bogen, Aleksandra Korolova, Alan Mislove, Aaron Rieke. 2019. Discrimination through optimization: How facebook's ad delivery can lead to biased outcomes. *Proceedings of the ACM on Human-Computer Interaction* **3**(CSCW) 1–30.

Alptekinoğlu, Aydın, Arunava Banerjee, Anand Paul, Nikhil Jain. 2013. Inventory pooling to deliver differentiated service. *Manufacturing & Service Operations Management* **15**(1) 33–44.

Atkinson, Anthony B. 1970. On the measurement of inequality. *Journal of economic theory* **2**(3) 244–263.

Balseiro, Santiago, Haihao Lu, Vahab Mirrokni. 2021. Regularized online allocation problems: Fairness and beyond. *International Conference on Machine Learning*. PMLR, 630–639.

Balseiro, Santiago R, Omar Besbes, Dana Pizarro. 2023. Survey of dynamic resource-constrained reward collection problems: Unified model and analysis. *Operations Research* .

Balseiro, Santiago R, Jon Feldman, Vahab Mirrokni, Shan Muthukrishnan. 2014. Yield optimization of display advertising with ad exchange. *Management Science* **60**(12) 2886–2907.

Barocas, Solon, Andrew D Selbst. 2016. Big data's disparate impact. *Calif. L. Rev.* **104** 671.

Bateni, MohammadHossein, Yiwei Chen, Dragos Florin Ciocan, Vahab Mirrokni. 2022. Fair resource allocation in a volatile marketplace. *Operations Research* **70**(1) 288–308.

Bernstein, Fernando, Sajad Modaresi, Denis Sauré. 2019. A dynamic clustering approach to data-driven assortment personalization. *Management Science* **65**(5) 2095–2115.

Bertsimas, Dimitris, Vivek F Farias, Nikolaos Trichakis. 2012. On the efficiency-fairness trade-off. *Management Science* **58**(12) 2234–2250.

Bumpensanti, Pornpawee, He Wang. 2020. A re-solving heuristic with uniformly bounded loss for network revenue management. *Management Science* **66**(7) 2993–3009.

Caro, Felipe, Victor Martínez-de Albéniz, Paat Rusmevichientong. 2014. The assortment packing problem: Multiperiod assortment planning for short-lived products. *Management Science* **60**(11) 2701–2721.

Celis, Elisa, Anay Mehrotra, Nisheeth Vishnoi. 2019. Toward controlling discrimination in online ad auctions. *International Conference on Machine Learning*. PMLR, 4456–4465.

Chen, Xi, Will Ma, David Simchi-Levi, Linwei Xin. 2023. Assortment planning for recommendations at checkout under inventory constraints. *Mathematics of Operations Research* .

Cheung, Wang Chi, David Simchi-Levi. 2017. Thompson sampling for online personalized assortment optimization problems with multinomial logit choice models. *Available at SSRN 3075658* .

Choi, Hana, Carl F Mela, Santiago R Balseiro, Adam Leary. 2020. Online display advertising markets: A literature review and future directions. *Information Systems Research* **31**(2) 556–575.

Dai, Jim G, Wuqin Lin. 2005. Maximum pressure policies in stochastic processing networks. *Operations Research* **53**(2) 197–218.

Davis, James, Guillermo Gallego, Huseyin Topaloglu. 2013. Assortment planning under the multinomial logit model with totally unimodular constraint structures. *Work in Progress* .

Davis, James M, Guillermo Gallego, Huseyin Topaloglu. 2014. Assortment optimization under variants of the nested logit model. *Operations Research* **62**(2) 250–273.

Dong, Lingxiu, Duo Shi, Fuqiang Zhang. 2022. 3D printing and product assortment strategy. *Management Science* **68**(8) 5724–5744.

Eppen, Gary D. 1979. Note—effects of centralization on expected costs in a multi-location newsboy problem. *Management science* **25**(5) 498–501.

Feldman, Jacob, Dennis J Zhang, Xiaofei Liu, Nannan Zhang. 2022. Customer choice models vs. machine learning: Finding optimal product displays on alibaba. *Operations Research* **70**(1) 309–328.

Feldman, Michael, Sorelle A Friedler, John Moeller, Carlos Scheidegger, Suresh Venkatasubramanian. 2015. Certifying and removing disparate impact. *proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*. 259–268.

Gallego, Guillermo, Anran Li, Van-Anh Truong, Xinshang Wang. 2016. Online personalized resource allocation with customer choice. Tech. rep., Working Paper. http://arxiv. org/abs/1511.01837 v1.

Gallego, Guillermo, Richard Ratliff, Sergey Shebalov. 2015. A general attraction model and sales-based linear program for network revenue management under customer choice. *Operations Research* **63**(1) 212–232.

Golrezaei, Negin, Hamid Nazerzadeh, Paat Rusmevichientong. 2014. Real-time optimization of personalized assortments. *Management Science* **60**(6) 1532–1551.

Hao, Xiaotian, Zhaoqing Peng, Yi Ma, Guan Wang, Junqi Jin, Jianye Hao, Shan Chen, Rongquan Bai, Mingzhou Xie, Miao Xu, Zhenzhe Zheng, Chuan Yu, HAN LI, Jian Xu, Kun Gai. 2020. Dynamic knapsack optimization towards efficient multi-channel sequential advertising. *ICML 2020* .

Hojjat, Ali, John Turner, Suleyman Cetintas, Jian Yang. 2017. A unified framework for the scheduling of guaranteed targeted display advertising under reach and frequency requirements. *Operations Research* **65**(2) 289–313.

Imana, Basileal, Aleksandra Korolova, John Heidemann. 2021. Auditing for discrimination in algorithms delivering job ads. *Proceedings of the Web Conference 2021*. 3767–3778.

Jasin, Stefanus, Sunil Kumar. 2012. A re-solving heuristic with bounded revenue loss for network revenue management with customer choice. *Mathematics of Operations Research* **37**(2) 313–345.

Jiang, Jiashuo, Shixin Wang, Jiawei Zhang. 2023. Achieving high individual service levels without safety stock? optimal rationing policy of pooled resources. *Operations Research* **71**(1) 358–377.

Kallus, Nathan, Madeleine Udell. 2020. Dynamic assortment personalization in high dimensions. *Operations Research* **68**(4) 1020–1037.

Kumar, Amit, Jon Kleinberg. 2000. Fairness measures for resource allocation. *Proceedings 41st annual symposium on foundations of computer science*. IEEE, 75–85.

Lambrecht, Anja, Catherine Tucker. 2019. Algorithmic bias? an empirical study of apparent gender-based discrimination in the display of stem career ads. *Management science* **65**(7) 2966–2981.

Lejeune, Miguel A, John Turner. 2019. Planning online advertising using gini indices. *Operations Research* **67**(5) 1222–1245.

Liu, Qian, Garrett Van Ryzin. 2008. On the choice-based linear programming model for network revenue management. *Manufacturing & Service Operations Management* **10**(2) 288–310.

Luce, R Duncan. 2012. *Individual choice behavior: A theoretical analysis*. Courier Corporation.

Lyu, Guodong, Wang-Chi Cheung, Mabel C Chou, Chung-Piaw Teo, Zhichao Zheng, Yuanguang Zhong. 2019. Capacity allocation in flexible production networks: Theory and applications. *Management Science* **65**(11) 5091–5109.

Lyu, Guodong, Mabel C Chou, Chung-Piaw Teo, Zhichao Zheng, Yuanguang Zhong. 2022. Stochastic knapsack revisited: The service level perspective. *Operations Research* **70**(2) 729–747.

Ma, Will, Pan Xu, Yifan Xu. 2020. Group-level fairness maximization in online bipartite matching. *arXiv preprint arXiv:2011.13908* .

McKeown, Nick, Adisak Mekkittikul, Venkat Anantharam, Jean Walrand. 1999. Achieving 100% throughput in an input-queued switch. *IEEE Transactions on Communications* **47**(8) 1260–1267.

Mehta, Aranyak, et al. 2013. Online matching and ad allocation. *Foundations and Trends® in Theoretical Computer Science* **8**(4) 265–368.

Microsoft. 2020. *Partner Incentives Co-op Guidebook: Business Policies for FY21*. Microsoft Corporation.

Mulvany, Justin, Ramandeep S Randhawa. 2021. Fair scheduling of heterogeneous customer populations. *Available at SSRN 3803016* .

Nakamura, Atsuyoshi, Naoki Abe. 2005. Improvements to the linear programming based scheduling of web advertisements. *Electronic Commerce Research* **5**(1) 75–98.

Nilforoshan, Hamed, Johann D Gaebler, Ravi Shroff, Sharad Goel. 2022. Causal conceptions of fairness and their consequences. *International Conference on Machine Learning*. PMLR, 16848–16887.

Rubin, Ronald B. 1978. The uniform guidelines on employee selection procedures: compromises and controversies. *Cath. UL Rev.* **28** 605.

Rusmevichientong, Paat, Zuo-Jun Max Shen, David B Shmoys. 2010. Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Operations Research* **58**(6) 1666–1680.

Shen, Huaxiao, Yanzhi Li, Youhua Chen, Kai Pan. 2021a. Integrated ad delivery planning for targeted display advertising. *Operations Research* **69**(5) 1409–1429.

Shen, Huaxiao, Yanzhi Li, Jingjing Guan, Geoffrey KF Tso. 2021b. A planning approach to revenue management for non-guaranteed targeted display advertising. *Production and Operations Management* **30**(6) 1583–1602.

Shi, Cong, Yehua Wei, Yuan Zhong. 2019. Process flexibility for multiperiod production systems. *Operations Research* **67**(5) 1300–1320.

Speicher, Till, Muhammad Ali, Giridhari Venkatadri, Filipe Nunes Ribeiro, George Arvanitakis, Fabrício Benevenuto, Krishna P Gummadi, Patrick Loiseau, Alan Mislove. 2018. Potential for discrimination in online targeted advertising. *Conference on Fairness, Accountability and Transparency*. PMLR, 5–19.

Stolyar, Alexander L. 2004. Maxweight scheduling in a generalized switch: State space collapse and workload minimization in heavy traffic. *The Annals of Applied Probability* **14**(1) 1–53.

Swaminathan, Jayashankar M, Ramesh Srinivasan. 1999. Managing individual customer service constraints under stochastic demand. *Operations Research Letters* **24**(3) 115–125.

Tassiulas, Leandros, Anthony Ephremides. 1990. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *29th IEEE Conference on Decision and Control*. IEEE, 2130–2132.

Turner, John. 2012. The planning of guaranteed targeted display advertising. *Operations Research* **60**(1) 18–33.

Wang, Ruxian. 2012. Capacitated assortment and price optimization under the multinomial logit model. *Operations Research Letters* **40**(6) 492–497.

Xu, Kuang, Yuan Zhong. 2020. Information and memory in dynamic resource allocation. *Operations Research* **68**(6) 1698–1715.

Yang, Jian, Erik Vee, Sergei Vassilvitskii, John Tomlin, Jayavel Shanmugasundaram, Tasos Anastasakos, Oliver Kennedy. 2012. Inventory allocation for online graphical display advertising using multi-objective optimization. *ICORES* **12** 293–304.

Ye, Zikun, Dennis J Zhang, Heng Zhang, Renyu Zhang, Xin Chen, Zhiwei Xu. 2022. Cold start to improve market thickness on online advertising platforms: Data-driven algorithms and field experiments. *Management Science* .

Young, H Peyton, R Mark Isaac. 1995. Equity: In theory and practice. *Journal of Economic Literature* **33**(1) 210–210.

Zhong, Yuanguang, Zhichao Zheng, Mabel C Chou, Chung-Piaw Teo. 2017. Resource pooling and allocation policies to deliver differentiated service. *Management Science* **64**(4) 1555–1573.

Online Appendices

A. Summary of Notations

Notation	Description
TTD	Two-stage target-debt framework
DWO	Debt-weighted offer-set policy
Fluid	Fluid-approximation approach
FV	Fairness-adjusted value
T	Total number of customers
$\mathcal{N}$	Set of ad campaigns
n	Number of ad campaigns
B_i	Total budget of ad campaign i
b_i	Bid price of ad campaign i per click-through
$j(t)$	Type of customer t
$\mathcal{M}$	Set of customer types
m	Number of customer types
p^j	Probability of a customer being type j
$S(t)$	Offer-set displayed to customer t
$\mathfrak{S}^j$	Collection of all possible offer-sets for type-j customers including ad targeting info
$y_i^j(t)$	Number of click-throughs by a type-j customer on ad i in time t
$\bar{y}_i^j$	Per-customer click-throughs of ad i by type-j customers
$\phi_i^j(S)$	Expected value of $y_i^j(t)$ conditioned on $S(t)=S$
$\mathsf{D}_{(j,y)}$	Joint customer type and click-through distribution
$\eta_i^{\mathcal{C}}$	Required click-throughs for customer-type set $\mathcal{C}$ on ad campaign i
$\mathfrak{K}_i$	Set of all $\mathcal{C}$ with $\eta_i^{\mathcal{C}}>0$ of ad i
r_i^j	Value of each click of ad campaign i by a type-j customer
$F(\boldsymbol{y})$	Fairness metric
$\mathcal{H}_{t-1}$	Realized history until the start of time t
Π	Set of policies
Π_{static}	Set of static policies
Π_{d}	Set of deterministic static policies
$(\mathcal{OP})$	Original stochastic program
$\mathcal{V}^*$	Optimal FV of the original stochastic program
α_j^i	Target for the per-period number of click-throughs of ad i from type-j customers
$\mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha})$	FV of the click-through target vector $\boldsymbol{\alpha}$
$(2\mathcal{SSP})$	Two-stage stochastic program
θ_i^j	Dual variable associated with satisfying the click-through target ad i from type-j customers
$(\mathcal{OTP})$	Reformulated optimal target problem
$\boldsymbol{\alpha}^*$	Solution to $(\mathcal{OTP})$
K	Maximum size of an offer-set
DWO-$\boldsymbol{\alpha}$	DWO policy with click-through target vector $\boldsymbol{\alpha}$
$d_i^j(t)$	Debt of the click-throughs from type-j customers on ad i in time t
$\mathcal{OP}(\gamma)$	Family of ad-allocation problems with scaling parameter γ
$\mathcal{V}(\pi\vert\gamma)$	Expected FV generated by policy π in $\mathcal{OP}(\gamma)$
$(\mathcal{OP}_{\mathsf{Fluid}})$	Fluid convex program
$\mathcal{V}_{\mathsf{Fluid}}(\boldsymbol{z})$	Objective value function of $(\mathcal{OP}_{\mathsf{Fluid}})$

Table 2 Summary of Notations

B. Soft-constraint Formulation

In this section, we consider the original stochastic program with soft constraints as follows:

$$\begin{aligned} \max_{\pi\in\Pi} \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^j y_i^j(t|\pi) + \lambda F(\bar{\boldsymbol{y}}(\pi)) - \nu \sum_{i\in\mathcal{N}}\sum_{\mathcal{C}\in\mathfrak{K}_i}\left(\frac{\eta_i^{\mathcal{C}}}{T} - \sum_{j\in\mathcal{C}}\bar{y}_i^j(\pi)\right)^+\right] \\ \text{s.t. } \frac{1}{T}\sum_{t=1}^{T}\sum_{j\in\mathcal{M}} b_i y_i^j(t|\pi) \le \frac{B_i}{T}, \ \textit{almost surely} \text{ for each } i\in\mathcal{N}, \end{aligned} \tag{$\mathcal{OP}_{Soft}$}$$

where the first term in the objective is the total per-customer-impression value from advertising, which we call the *efficiency* of policy π denoted by $\mathcal{E}(\pi) := \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^j y_i^j(t|\pi)\right]$, the second term is the *fairness* of policy π denoted by $\lambda\cdot\mathcal{F}(\pi) := \lambda\cdot\mathbb{E}[F(\bar{\boldsymbol{y}}(\pi))]$, and the third term in the objective is the soft constraints of the click-through requirements denoted by $\nu\cdot\mathcal{G}(\pi) := -\nu\cdot\mathbb{E}\left[\sum_{i\in\mathcal{N}}\sum_{\mathcal{C}\in\mathfrak{K}_i}\left(\frac{\eta_i^{\mathcal{C}}}{T} - \sum_{j\in\mathcal{C}}\bar{y}_i^j(\pi)\right)^+\right]$ with a parameter $\nu > 0$. Hence, the total objective value denoted by FV_{Soft} under policy π is given by $\mathcal{V}_{Soft}(\pi) := \mathcal{E}(\pi) + \lambda\cdot\mathcal{F}(\pi) + \nu\cdot\mathcal{G}(\pi)$, and we denote the optimal FV_{Soft} as $\mathcal{V}^*_{Soft} = \limsup_{\pi\in\Pi}\mathcal{V}_{Soft}(\pi)$ and the optimal policy (if it exists) as $\pi^* = \arg\max_{\pi\in\Pi}\mathcal{V}_{Soft}(\pi)$. We also remark that the constraint of ($\mathcal{OP}_{Soft}$) refers to the budget constraint of each ad.

We formulate the original stochastic program ($\mathcal{OP}_{Soft}$) as a dynamic program (DP). Specifically, we define $Y_i^j(t) := \sum_{\tau=1}^{t-1} y_i^j(\tau)$ as the accumulative number of click-throughs until the start of time t, and

$$\begin{aligned} \mathbb{V}_t(\boldsymbol{Y}(t)) := \max_{\pi\in\Pi}\mathbb{E}\left[\sum_{\tau=t}^{T}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^j y_i^j(t|\pi) + T\lambda F(\bar{\mathbf{y}}(\pi)) - \nu\sum_{i\in\mathcal{N}}\sum_{\mathcal{C}\in\mathfrak{K}_i}\left(\eta_i^{\mathcal{C}} - \sum_{t=1}^{T}\sum_{j\in\mathcal{C}} y_i^j(t)\right)^+ \middle| \boldsymbol{Y}(t)\right] \\ \text{s.t. } \sum_{\tau=t}^{T}\sum_{j\in\mathcal{M}} b_i y_i^j(t|\pi) \le B_i - \sum_{j\in\mathcal{M}} b_i Y_i^j(t), \ \textit{almost surely} \text{ for each } i\in\mathcal{N}. \end{aligned} \tag{19}$$

Hence, $\mathbb{V}_t(\boldsymbol{Y}(t))$ is the maximum expected FV_{Soft} given that the number of accumulative click-throughs at the beginning of time t is $\boldsymbol{Y}(t) := (Y_i^j(t) : i\in\mathcal{N}, j\in\mathcal{M})$.

To formulate the DP, we first specify the boundary/terminal value function $\mathbb{V}_{T+1}(\boldsymbol{Y}(T+1))$. To this end, we define

$$\mathcal{Y} := \left\{\boldsymbol{Y}(T+1)\in\mathbb{R}_+^{nm} : b_i\sum_{j\in\mathcal{M}} Y_i^j(T+1) \le B_i \text{ for each } i\in\mathcal{N}\right\}$$

as the feasible region for the accumulative number of click-throughs for the entire planning horizon, $\boldsymbol{Y}(T+1)$. The boundary value function is defined as follows:

$$\mathbb{V}_{T+1}(\boldsymbol{Y}(T+1)) = \begin{cases} T\lambda F\left(\frac{\boldsymbol{Y}(T+1)}{T}\right) - \nu\sum_{i\in\mathcal{N}}\sum_{\mathcal{C}\in\mathfrak{K}_i}\left(\eta_i^{\mathcal{C}} - \sum_{j\in\mathcal{C}} Y_i^j(T+1)\right)^+, & \text{if } \boldsymbol{Y}(T+1)\in\mathcal{Y}, \\ -\bar{M}, & \text{otherwise}, \end{cases} \tag{20}$$

where $\bar{M}$ is a sufficiently large positive number that is far bigger than $\mathcal{V}^*_{Soft}$ (e.g., $\bar{M} := C\cdot\max\{\mathcal{V}^*_{Soft}, 1\}$, where $C > 0$ is a very large positive number).

By the standard backward induction argument, we are now ready to write the Bellman equation to evaluate $\mathbb{V}_t(\boldsymbol{Y}(t))$ in (19):

$$\mathbb{V}_t(\boldsymbol{Y}(t)) = \sum_{j\in\mathcal{M}} p^j \max_{S(t)\in\mathfrak{S}^j}\mathbb{E}_{\boldsymbol{y}(t)}\left[\sum_{i\in\mathcal{N}} r_i^j y_i^j + \mathbb{V}_{t+1}(\boldsymbol{Y}(t)+\boldsymbol{y}(t))\middle| S(t), j(t) = j\right]. \tag{21}$$

Therefore, the optimal FV$_{Soft}$ for the original problem ($\mathcal{OP}_{Soft}$) is

$$\mathcal{V}^*_{Soft} = \frac{\mathbb{V}_1(\mathbf{0})}{T}, \text{ where } \mathbf{0} := (0,0,...,0)' \in \mathbb{R}^{nm}.$$

Due to the curse of dimensionality, the above DP formulation of ($\mathcal{OP}_{Soft}$) is intractable even when m and n are just moderately large. Therefore, we resort to our TTD framework and the induced DWO algorithm to solve the ad-allocation optimization problem. We can construct a related optimal target problem as follows:

$$\begin{aligned}
&\max_{\boldsymbol{\alpha} \geq \mathbf{0}} \ \mathcal{V}_{Soft-\mathsf{CT}}(\boldsymbol{\alpha}) \\
&\text{s.t. } h(\boldsymbol{\alpha}) \geq 0, \\
&\quad b_i \sum_{j\in\mathcal{M}} \alpha_i^j \leq \frac{B_i}{T}, \text{ for each } i \in \mathcal{N},
\end{aligned} \tag{$\mathcal{OTP}_{Soft}$}$$

where $\mathcal{V}_{Soft-\mathsf{CT}}(\boldsymbol{\alpha}) := \sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{M}} r_i^j \alpha_i^j + \lambda F(\boldsymbol{\alpha}) + \nu G(\boldsymbol{\alpha})$ and $G(\boldsymbol{\alpha}) := -\sum_{i\in\mathcal{N}} \sum_{\mathcal{C}\in\mathfrak{K}_i} \left(\frac{\eta_i^{\mathcal{C}}}{T} - \sum_{j\in\mathcal{C}} \alpha_i^j\right)^+$. It is straightforward to check that $\mathcal{V}_{Soft-\mathsf{CT}}(\cdot)$ is concave in $\boldsymbol{\alpha}$ and the constraint $h(\boldsymbol{\alpha}) \geq 0$ is equivalent to (3). Solving the convex program ($\mathcal{OTP}_{Soft}$) obtains the optimal targets $\boldsymbol{\alpha}^*_{Soft}$. Then, we apply the DWO-$\boldsymbol{\alpha}^*_{Soft}$ policy to dynamically display the offer-set to each arriving customer. Following the same analysis as our main model with the expected click-through requirements, we can show that the DWO-$\boldsymbol{\alpha}^*_{Soft}$ policy is also asymptotically optimal for ($\mathcal{OP}_{Soft}$). To avoid repetition, we omit the proof details.

C. Metrics of Fairness

In this section, we describe a few commonly adopted metrics of fairness $F(\cdot)$, all of which are concave and can be coherently incorporated into our framework.

Max-min fairness. The recent trend of machine-learning fairness has promoted that minority customers should have sufficient click-throughs in a recommender/advertising system; otherwise, their needs cannot be well taken care of due to data scarcity. A natural choice to accommodate such fairness concern is the max-min fairness metric, which has been extensively studied in the literature of economics (e.g., Young and Isaac 1995), computer science (e.g., Kumar and Kleinberg 2000), and operations research (e.g., Bertsimas et al. 2012). Specifically, we define function $F(\cdot)$ as follows:

$$F(\bar{\boldsymbol{y}}) = \min_{j\in\mathcal{M}} \left\{ \sum_{i\in\mathcal{N}} \bar{y}_i^j \right\}. \tag{22}$$

Max-min fairness drives the platform to maximize the minimum per-customer-impression click-throughs from all customer types, ensuring that no customer type has too few click-throughs. We also note that max-min fairness can also be evaluated with respect to advertisers, i.e., $F(\bar{\boldsymbol{y}}) = \min_{i\in\mathcal{N}} \left\{ \sum_{j\in\mathcal{M}} \bar{y}_i^j \right\}$, so as to ensure that no advertiser receives too few click-throughs. One may also generalize max-min fairness to the $\alpha-$fairness metric (see, e.g., Bertsimas et al. 2012), i.e., $F(\bar{\boldsymbol{y}}) = \sum_{j\in\mathcal{M}} \frac{1}{1-\alpha} \left(\sum_{i\in\mathcal{N}} \bar{y}_i^j \right)^{1-\alpha}$ if $\alpha \neq 1$ and $F(\bar{\boldsymbol{y}}) = \sum_{j\in\mathcal{M}} \log \left(\sum_{i\in\mathcal{N}} \bar{y}_i^j \right)$ if $\alpha = 1$, which is reduced to the max-min fairness metric if we take $\alpha \to +\infty$.

Gini mean difference fairness. Advertisers generally prefer receiving impressions/click-throughs/conversions that are evenly spread across their targeted customer types (e.g., Lejeune and Turner 2019). One way to capture such preference is through Gini mean difference (GMD) fairness. The Gini coefficient/index has long been a canonical measure of income inequality in economics (e.g., Atkinson 1970), and it has recently been studied in the advertising literature to maximize the spreading of impressions across targeted user types (e.g., Lejeune and Turner 2019). Following Lejeune and Turner (2019), given the average click-through vector $\bar{\boldsymbol{y}}_i = (\bar{y}_i^1, \bar{y}_i^2, ..., \bar{y}_i^m)'$ of ad i, we first define the GMD fairness for each ad i:

$$GMD_i(\bar{\boldsymbol{y}}_i) = \frac{2}{(\sum_{j\in\mathcal{M}} p^j)^2} \sum_{j,j'\in\mathcal{M}} p^j p^{j'} \left| \frac{\bar{y}_i^j}{p^j} - \frac{\bar{y}_i^{j'}}{p^{j'}} \right|, \tag{23}$$

where p^j is the proportion of type-j customers, and $\frac{\bar{y}_i^j}{p^j} = \frac{\sum_{t=1}^T y_i^j(t)}{p^j T}$ is the per-type-j customer click-throughs of ad i. Hence, the Gini coefficient of ad i is defined as follows:

$$G_i(\bar{\boldsymbol{y}}_i) = \frac{\left(\sum_{j\in\mathcal{M}} p^j\right) GMD_i(\bar{\boldsymbol{y}}_i)}{2\sum_{j\in\mathcal{M}} \bar{y}_i^j}.$$

We are now ready to define the GMD fairness as a weighted sum of the negative Gini coefficient of each ad:

$$F(\bar{\boldsymbol{y}}) = -\sum_{i\in\mathcal{N}} k_i G_i(\bar{\boldsymbol{y}}_i) = -\sum_{i\in\mathcal{N}} \frac{k_i}{\left(\sum_{j\in\mathcal{M}} \bar{y}_i^j\right)\left(\sum_{j\in\mathcal{M}} p^j\right)} \sum_{j,j'\in\mathcal{M}} |p^{j'}\bar{y}_i^j - p^j \bar{y}_i^{j'}|,$$

where $k_i \geq 0$ is the weight of ad i according to its importance in the GMD fairness metric. Following Lejeune and Turner (2019), we choose $k_i = \sum_{j\in\mathcal{M}} \bar{y}_i^j$ which gives rise to our GMD fairness metric as (24):

$$F(\bar{\boldsymbol{y}}) = -\sum_{i\in\mathcal{N}} \frac{1}{\sum_{j\in\mathcal{M}} p^j} \sum_{j,j'\in\mathcal{M}} |p^{j'}\bar{y}_i^j - p^j \bar{y}_i^{j'}|. \tag{24}$$

It is clear from (24) that the GMD fairness metric prompts the platform to induce click-throughs from each targeted customer type j proportional to its traffic p^j.

Disparate impact. Disparate impact is a widely discussed algorithmic discrimination measure (e.g. Feldman et al. 2015, Barocas and Selbst 2016). We show that this measure can be adopted in $F(\cdot)$. Following Feldman et al. (2015), we call there exists disparate impact for type-j customers if

$$\frac{\bar{y}_i^j/p^j}{\bar{y}_i^{j'}/p^{j'}} \leq \tau \text{ for some } i \text{ and } j' \neq j,$$

where τ is a parameter in $[0,1]$. In practice, τ is usually set as 0.8 given the prevailing $80\% - 20\%$ rule (see, e.g. Rubin 1978). Hence, we can define the fairness metric $F(\cdot)$ for eliminating disparate impact as follows:

$$F(\bar{\boldsymbol{y}}) = \min_{i\in\mathcal{N}, j\in\mathcal{J}_i} \left\{ \frac{\bar{y}_i^j}{p^j} - \tau \cdot \max_{j'\in\mathcal{J}_i'} \frac{\bar{y}_i^{j'}}{p^{j'}} \right\}, \tag{25}$$

where $\mathcal{J}_i$ and $\mathcal{J}_i'$ ($\mathcal{J}_i, \mathcal{J}_i' \in \mathcal{M}$) denote the sets of minority and majority types, respectively. It is straightforward to check that (25) is concave, which measures the disparate impact attributed to the ad allocation algorithm.

D. Proof of Statements

We provide the proof of all the technical results in this section.

Proof of Lemma 1

It is evident that any feasible click-through target vector $\boldsymbol{\alpha}$ to (2$\mathcal{SSP}$) must also be feasible to (2), because the feasible region of (2$\mathcal{SSP}$) is a subset of that of (2).

We now show that any feasible click-through target vector $\boldsymbol{\alpha}$ to (2) is also feasible to (2$\mathcal{SSP}$). Consider a feasible $\boldsymbol{\alpha}$ to (2). We have, there exists a feasible policy $\pi \in \Pi$,

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} y_i^j(t|\pi)\right] \geq \alpha_i^j, \text{ for each } i \in \mathcal{N}, j \in \mathcal{M}.$$

We define the following probability measure induced by π, $z_{\mathsf{static}}(S|\pi)$: For any $j \in \mathcal{M}$ and $S \in \mathfrak{S}^j$,

$$z^j_{\mathsf{static}}(S|\pi) := \mathbb{P}[\pi_{\tilde{t}}(j(\tilde{t}), \mathcal{H}_{\tilde{t}-1}) = S | j(\tilde{t}) = j],$$

where $\tilde{t}$ is a random variable uniformly distributed on $\{1, 2, ..., T\}$ and independent of everything else. Based on $\boldsymbol{z}_{\mathsf{static}}(\cdot|\pi)$, we construct a static policy π_{static}, which selects offer-set S given each customer type j with probability $z^j_{\mathsf{static}}(S|\pi)$. Straightforward algebraic manipulations and the law of iterated expectations together yield that

$$\mathbb{E}\left[y_i^j(t|\pi_{\mathsf{static}})\right] = \sum_{S \in \mathfrak{S}^j} p^j \phi(j|S) z^j_{\mathsf{static}}(S|\pi) = \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} y_i^j(t|\pi)\right] \geq \alpha_i^j, , \text{ for each } i \in \mathcal{N}, j \in \mathcal{M}.$$

Hence, π_{static} and $\boldsymbol{\alpha}$ are feasible to (2$\mathcal{SSP}$). Moreover, because these two problems have the same objective function, any optimal click-through target vector for one problem must be also optimal for another. This completes the proof of Lemma 1. □

Proof of Theorem 1

The proof follows from the discussions before (8). □

Proof of Proposition 1

Before proving Proposition 1, we first state and prove a few auxiliary results. It is sometimes more convenient to use a binary variable representation of a deterministic static offer-set policy $\pi \in \Pi_{\mathsf{d}}$. More specifically, $\pi \in \Pi_{\mathsf{d}}$ can be equivalently represented by an nm-dimensional binary vector $\boldsymbol{x} = (x_i^j \in \{0,1\} : i \in \mathcal{N}, j \in \mathcal{M})$, where $x_i^j = 1$ means that $i \in \pi(j)$, i.e., ad i is included in the offer-set displayed to a type-j customer. With a slight abuse of notation, we denote $\phi_i^j(\boldsymbol{x})$ as the expected click-throughs of a type-j customer for ad i if the offer-set displayed to this customer is $S^j = \{i \in \mathcal{N} : x_i^j = 1\}$. Under the MNL model, we have

$$\phi_i^j(\boldsymbol{x}) = \frac{v_i^j x_i^j}{1 + \sum_{i' \in \mathcal{N}} v_{i'}^j x_{i'}^j}, \tag{26}$$

where $v_i^j > 0$ is the attractiveness of ad i to type-j customers (see, also, (11)). Denote the set of all plausible offer-set representation vectors as $\mathcal{X} \subset \{0,1\}^{nm}$, and the set of plausible offer-set representation vectors displayed to a type-j customer as $\mathcal{X}^j \subset \{0,1\}^n$. Applying Theorems 1 to the MNL choice model (26), we have the following corollary.

Corollary 1. *If customers follow the MNL click-through model (26), a click-through target vector $\boldsymbol{\alpha}$ is single-period feasible if and only if*

$$\max_{\boldsymbol{x}\in\mathcal{X}} \sum_{i\in\mathcal{N},j\in\mathcal{M}} \frac{p^j v_i^j \theta_i^j x_i^j}{1+\sum\limits_{i'\in\mathcal{N}} v_{i'}^j x_{i'}^j} \geq \sum_{i\in\mathcal{N},j\in\mathcal{M}} \alpha_i^j \theta_i^j \text{ for all } \theta_i^j \geq 0 \ (i\in\mathcal{N}, j\in\mathcal{M}). \tag{27}$$

Furthermore, (27) is equivalent to, for each $j\in\mathcal{M}$

$$\max_{\boldsymbol{x}^j\in\mathcal{X}^j} \sum_{i\in\mathcal{N}} \frac{p^j v_i^j \theta_i^j x_i^j}{1+\sum\limits_{i'\in\mathcal{N}} v_{i'}^j x_{i'}^j} \geq \sum_{i\in\mathcal{N}} \alpha_i^j \theta_i^j \text{ for all } \theta_i^j \geq 0 \ (i\in\mathcal{N}). \tag{28}$$

Proof of Corollary 1

Directly applying Theorem 1 to the MNL choice model implies that $\boldsymbol{\alpha}$ is feasible if and only if inequality (27) holds.

We now show that (27) implies (28). If (27) holds for any $\boldsymbol{\theta}\geq\mathbf{0}$, then it also holds for any $\boldsymbol{\theta}\geq\mathbf{0}$ with $\theta_i^{j'}=0$ (for $j'\neq j$ and all $i\in\mathcal{N}$). Therefore, (28) holds.

Finally, we show that if (28) holds for all $j\in\mathcal{M}$, (27) holds as well. Note that the left-hand side of (27) can be decomposed into independent parts as follows:

$$\max_{\boldsymbol{x}\in\mathcal{X}} \sum_{i\in\mathcal{N},j\in\mathcal{M}} \frac{p^j v_i^j \theta_i^j x_i^j}{1+\sum\limits_{i'\in\mathcal{N}} v_{i'}^j x_{i'}^j} = \max_{\boldsymbol{x}\in\mathcal{X}} \sum_{j\in\mathcal{M}} \sum_{i\in\mathcal{N}} \frac{p^j v_i^j \theta_i^j x_i^j}{1+\sum\limits_{i'\in\mathcal{N}} v_{i'}^j x_{i'}^j} = \sum_{j\in\mathcal{M}} \max_{\boldsymbol{x}^j\in\mathcal{X}^j} \sum_{i\in\mathcal{N}} \frac{p^j v_i^j \theta_i^j x_i^j}{1+\sum\limits_{i'\in\mathcal{N}} v_{i'}^j x_{i'}^j} \geq \sum_{j\in\mathcal{M}} \sum_{i\in\mathcal{N}} \alpha_i^j \theta_i^j = \sum_{i\in\mathcal{N},j\in\mathcal{M}} \alpha_i^j \theta_i^j,$$

for any $\boldsymbol{\theta}\geq\mathbf{0}$, where the inequality follows from (28). Therefore, that (27) holds is equivalent to that (28) holds for all $j\in\mathcal{M}$. This completes the proof of Corollary 1. $\square$

Leveraging the structural properties of the MNL model, we can give a sharper and simpler characterization for the feasibility condition (as the solution to a linear program). The following lemma characterizes the feasibility condition for a click-through target vector $\boldsymbol{\alpha}$, taking into account the cardinality constraint that the size of an offer-set displayed to any customer is upper bounded by K, i.e., $|S(t)|\leq K$ for any customer t.

Lemma 3. *If customers follow the MNL click-through model (26) and the set of all feasible offer-sets is $\mathfrak{S}^j=\{S\subset\mathcal{N}:|S|\leq K\}$ for each $j\in\mathcal{M}$, we have $\boldsymbol{\alpha}$ is single-period feasible if and only if there exist $\boldsymbol{w}:=(w_i^j: i\in\mathcal{N}, j\in\mathcal{M})$ and $\boldsymbol{z}:=(z^j: j\in\mathcal{M})$ that satisfy the following linear constraints*

$$\begin{aligned} &p^j v_i^j w_i^j \geq \alpha_i^j,\ w_i^j\leq z^j,\ w_i^j\geq 0, \text{ for each } i\in\mathcal{N}, j\in\mathcal{M},\\ &\sum_{i\in\mathcal{N}} v_i^j w_i^j + z^j = 1,\ \sum_{i\in\mathcal{N}} w_i^j \leq K z^j, \text{ for each } j\in\mathcal{M}, \end{aligned} \tag{29}$$

where $z^j := \frac{1}{1+\sum\limits_{i'\in\mathcal{N}} v_{i'}^j x_{i'}^j}$ *and* $w_i^j := x_i^j z^j = \frac{x_i^j}{1+\sum\limits_{i'\in\mathcal{N}} v_{i'}^j x_{i'}^j}$.

Proof of Lemma 3

A standard result in fractional programming postulates that the left-hand side of (28) is quasi-convex in $\boldsymbol{x}^j$ for all $j\in\mathcal{M}$, so there always exists a maximizer on the boundary of the feasible region. Thus, we can relax the binary constraint $x_i^j\in\{0,1\}$ to $x_i^j\in[0,1]$ in (28), which is therefore equivalent to

$$\max_{\boldsymbol{x}^j\in[0,1]^n,\ \sum\limits_{i\in\mathcal{N}} x_i^j\leq K} \sum_{i\in\mathcal{N}} \frac{p^j v_i^j \theta_i^j x_i^j}{1+\sum_{i'\in\mathcal{N}} v_{i'}^j x_{i'}^j} \geq \sum_{i\in\mathcal{N}} \alpha_i^j \theta_i^j \text{ for all } \boldsymbol{\theta}^j\geq\mathbf{0} \text{ and } j\in\mathcal{M}. \tag{30}$$

We change the decision variable and define, for all $j \in \mathcal{M}$,

$$z^j := \frac{1}{1 + \sum_{i' \in \mathcal{N}} v_{i'}^j x_{i'}^j} \text{ and } w_i^j := x_i^j z^j = \frac{x_i^j}{1 + \sum_{i' \in \mathcal{N}} v_{i'}^j x_{i'}^j}.$$

Then, we can rewrite (30) as, for any $j \in \mathcal{M}$,

$$\begin{aligned} &\min_{\boldsymbol{\theta}^j \geq \mathbf{0}} \left(\max_{w_i^j, z^j} \sum_{i \in \mathcal{N}} p^j v_i^j w_i^j \theta_i^j - \sum_{i \in \mathcal{N}} \alpha_i^j \theta_i^j \right) \geq 0 \\ &\text{s.t. } \sum_{i \in \mathcal{N}} v_i^j w_i^j + z^j = 1, \\ &\qquad \sum_{i \in \mathcal{N}} w_i^j \leq K z^j, \\ &\qquad 0 \leq w_i^j \leq z^j, \text{ for each } i \in \mathcal{N}. \end{aligned} \tag{31}$$

By Sion's minimax theorem, we can exchange the maximization and minimization operators so that (31) is equivalent to, for any $j \in \mathcal{M}$:

$$\begin{aligned} \max_{\boldsymbol{w}^j, z^j} \min_{\boldsymbol{\theta}^j \geq 0} &\sum_{i \in \mathcal{N}} \theta_i^j (p^j v_i^j w_i^j - \alpha_i^j) \geq 0, \\ \text{s.t. } &\sum_{i \in \mathcal{N}} v_i^j w_i^j + z^j = 1, \\ &\sum_{i \in \mathcal{N}} w_i^j \leq K z^j, \\ &0 \leq w_i^j \leq z^j, \text{ for each } i \in \mathcal{N}. \end{aligned} \tag{32}$$

Therefore, (32) holds if and only if there exist $\boldsymbol{w}^j$ and z^j such that all the constraints in (32) hold and $\sum_{i \in \mathcal{N}} \theta_i^j (p^j v_i^j w_i^j - \alpha_i^j) \geq 0$ holds for all $\boldsymbol{\theta}^j \geq \mathbf{0}$, which is equivalent to $p^j v_i^j w_i^j - \alpha_i^j \geq 0$ for all $i \in \mathcal{N}$. Therefore, (32) is equivalent to that, for any $j \in \mathcal{M}$,

$$\begin{aligned} &p^j v_i^j w_i^j - \alpha_i^j \geq 0, \text{ for each } i \in \mathcal{N}, \\ &\sum_{i \in \mathcal{N}} v_i^j w_i^j + z^j = 1, \\ &\sum_{i \in \mathcal{N}} w_i^j \leq K z^j, \\ &0 \leq w_i^j \leq z^j, \text{ for each } i \in \mathcal{N}. \end{aligned} \tag{33}$$

That (33) holds for all $j \in \mathcal{M}$ is equivalent to that (29) holds. This completes the proof of Lemma 3. $\square$

We now prove Proposition 1 itself. It suffices to show that, taking into account the cardinality constraint $|S| \leq K$, the (first-stage) feasible region for the first-stage click-through target vector $\boldsymbol{\alpha}$ is given by the following linear constraints:

$$\mathcal{A}_{MNL} := \left\{ \boldsymbol{\alpha} \in \mathbb{R}_+^{nm} : \sum_{i' \in \mathcal{N}} \alpha_{i'}^j + \frac{\alpha_i^j}{v_i^j} \leq p^j, \text{ for each } i \in \mathcal{N}, j \in \mathcal{M}, \text{ and } \sum_{i \in \mathcal{N}} \alpha_i^j + \frac{1}{K} \sum_{i \in \mathcal{N}} \frac{\alpha_i^j}{v_i^j} \leq p^j, \text{ for each } j \in \mathcal{M} \right\}. \tag{34}$$

We first show that if (29) holds, then $\boldsymbol{\alpha} \in \mathcal{A}_{MNL}$. By the first inequality of (29), we have $v_i^j w_i^j \geq \frac{\alpha_i^j}{p^j}$ for all $i \in \mathcal{N}$ and $j \in \mathcal{M}$. Plugging this into the first equality of (29), we have

$$1 - z^j = \sum_{i \in \mathcal{N}} v_i^j w_i^j \geq \sum_{i \in \mathcal{N}} \frac{\alpha_i^j}{p^j}, \text{ for each } j \in \mathcal{M}.$$

Thus, by the first and second inequalities of (29), we have

$$\sum_{i' \in \mathcal{N}} \frac{\alpha_{i'}^j}{p^j} \le 1 - z^j \le 1 - w_i^j \le 1 - \frac{\alpha_i^j}{p^j v_i^j} \text{ for each } i \in \mathcal{N}, j \in \mathcal{M}.$$

Rearranging the terms, we have

$$p^j \ge \sum_{i' \in \mathcal{N}} \alpha_{i'}^j + \frac{\alpha_i^j}{v_i^j} \text{ for each } i \in \mathcal{N}, j \in \mathcal{M}.$$

The first, second, and fourth inequalities and the first equality of (29) imply that

$$\sum_{i \in \mathcal{N}} \frac{\alpha_i^j}{p^j v_i^j} \le \sum_{i \in \mathcal{N}} w_i^j \le K z^j = K(1 - \sum_{i \in \mathcal{N}} v_i^j w_i^j) \le K \left(1 - \sum_{i \in \mathcal{N}} \frac{\alpha_i^j}{p^j}\right) \text{ for each } j \in \mathcal{M}.$$

Rearranging the terms, we have

$$p^j \ge \sum_{i \in \mathcal{N}} \alpha_i^j + \frac{1}{K} \sum_{i \in \mathcal{N}} \frac{\alpha_i^j}{v_i^j} \text{ for each } j \in \mathcal{M}.$$

Therefore, if (29) holds, we have $\boldsymbol{\alpha} \in \mathcal{A}_{MNL}$.

Next, we show that if $\boldsymbol{\alpha} \in \mathcal{A}_{MNL}$, (29) holds. Given $\boldsymbol{\alpha} \in \mathcal{A}_{MNL}$, define

$$w_i^j = \frac{\alpha_i^j}{p^j v_i^j} \text{ for each } i \in \mathcal{N} \text{ and } j \in \mathcal{M}, \text{ and } z^j = 1 - \sum_{i \in \mathcal{N}} \frac{\alpha_i^j}{p^j} \text{ for each } j \in \mathcal{M}.$$

To show (29), it suffices to show the first, second and fourth inequalities hold because the rest of the constraints hold trivially.

Since $p^j \ge \sum_{i \in \mathcal{N}} \alpha_i^j + \frac{1}{K} \sum_{i \in \mathcal{N}} \frac{\alpha_i^j}{v_i^j}$ for each $j \in \mathcal{M}$, we have

$$\sum_{i \in \mathcal{N}} w_i^j = \sum_{i \in \mathcal{N}} \frac{\alpha_i^j}{p^j v_i^j} = \frac{1}{p^j} \sum_{i \in \mathcal{N}} \frac{\alpha_i^j}{v_i^j} \le K \left(1 - \sum_{i \in \mathcal{N}} \frac{\alpha_i^j}{p^j}\right) = K z^j \text{ for each } j \in \mathcal{M}.$$

Hence, the second inequality of (29) holds. Since $p^j \ge \sum_{i' \in \mathcal{N}} \alpha_{i'}^j + \frac{\alpha_i^j}{v_i^j}$ for each $i \in \mathcal{N}, j \in \mathcal{M}$, we have

$$w_i^j = \frac{\alpha_i^j}{p^j v_i^j} \le 1 - \sum_{i' \in \mathcal{N}} \frac{\alpha_{i'}^j}{p^j} = z^j \text{ for each } i \in \mathcal{N}, j \in \mathcal{M}.$$

Therefore, (29) holds. Hence, the first-stage feasible region of $\boldsymbol{\alpha}$ is characterized by (34). This completes the proof of Proposition 1. $\square$

Proof of Theorem 2

Let us consider a problem identical to $\mathcal{OP}(\gamma)$ but without budget constraints (i.e., $B_i(\gamma) = +\infty$ for all $i \in \mathcal{N}$ and $\gamma > 0$), which we denote as $\mathcal{OP}_*(\gamma)$. By definition, in $\mathcal{OP}_*(\gamma)$, any ad i will not run out of budget throughout the planning horizon. Throughout the proof of Theorem 2, we write $y_i^j(t) = y_i^j(t|\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}))$ whenever there is no confusion.

- *Step 1.* For problem $\mathcal{OP}_*(\gamma)$, if $\boldsymbol{\alpha}$ is single-period feasible, it holds that

$$\liminf_{\gamma \uparrow +\infty} \frac{1}{T(\gamma)} \sum_{t=1}^{T(\gamma)} y_i^j(t) \ge \alpha_i^j \text{ for all } i \in \mathcal{N} \text{ and } j \in \mathcal{M}. \tag{35}$$

Under the $\mathsf{DWO}\text{-}\boldsymbol{\alpha}$ algorithm, we have that

$$t\alpha_i^j - \sum_{\tau=1}^{t} y_i^j(\tau) = d_i^j(t+1) \le \left(d_i^j(t+1)\right)^+.$$

Therefore, it suffices to show that, if (8) holds,

$$\limsup_{t\to+\infty} \frac{1}{t}\cdot d_i^j(t+1) \le 0, \text{ almost surely in problem } \mathcal{OP}_*(+\infty).$$

For a vector $\boldsymbol{x}\in\mathbb{R}^n$, we use $\boldsymbol{x}^+$ to denote the component-wise positive part of $\boldsymbol{x}$. Also note that, for any $A, B\in\mathbb{R}$, $((A+B)^+)^2 \le (A^+ + B)^2$. We have

$$\begin{aligned}\mathbb{E}||\left(\boldsymbol{d}(t+1)\right)^+||_2^2 =& \mathbb{E}||\left(\boldsymbol{d}(t)+\boldsymbol{\alpha}-\boldsymbol{y}(t)\right)^+||_2^2 \le \mathbb{E}||(\boldsymbol{d}(t))^+ + \boldsymbol{\alpha}-\boldsymbol{y}(t)||_2^2\\ =& \mathbb{E}||(\boldsymbol{d}(t))^+||_2^2 + \mathbb{E}||\boldsymbol{\alpha}-\boldsymbol{y}(t)||_2^2 + 2\mathbb{E}\left[\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}}(d_i^j(t))^+\cdot\alpha_i^j - \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}}(d_i^j(t))^+\cdot y_i^j(t)\right],\end{aligned}$$

where $||\cdot||_2$ denotes the ℓ_2-norm in a Euclidean space. Since $(d_i^j(t))^+ \ge 0$ for all $i\in\mathcal{N}$ and $j\in\mathcal{M}$, inequality (8) implies that

$$\mathbb{E}\left[\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}}(d_i^j(t))^+\cdot\alpha_i^j - \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}}(d_i^j(t))^+\cdot y_i^j(t)\right] \le 0.$$

Furthermore,

$$\mathbb{E}||\boldsymbol{\alpha}-\boldsymbol{y}(t)||_2^2 \le n\cdot m\cdot \max_{i\in\mathcal{N},j\in\mathcal{M}}\left\{\mathbb{E}[(y_i^j(t))^2]+(\alpha_i^j)^2\right\} \le n\cdot m\cdot C, \text{ where } C := \max_{i\in\mathcal{N},j\in\mathcal{M}}(\alpha_i^j)^2+1\le 2.$$

Therefore,

$$\mathbb{E}||(\boldsymbol{d}(t+1))^+||_2^2 \le ||(\boldsymbol{d}(1))^+||_2^2 + tnmC \text{ for all } t\ge 1. \tag{36}$$

By Jensen's inequality and that $||\cdot||_2^2$ is convex,

$$||\mathbb{E}[(\boldsymbol{d}(t+1))^+]||_2^2 \le \mathbb{E}||(\boldsymbol{d}(t+1))^+||_2^2 \le tnmC \text{ for all } t\ge 1. \tag{37}$$

Therefore,

$$0\le \frac{1}{t}||\mathbb{E}[(\boldsymbol{d}(t+1))^+]||_2 \le \sqrt{\frac{nmC}{t}}, \text{ which implies that } \limsup_{t\to+\infty}\frac{1}{t}||\mathbb{E}[(\boldsymbol{d}(t+1))^+]||_2 = 0.$$

Hence

$$\limsup_{t\to+\infty}\frac{1}{t}\cdot(d_i^j(t+1))^+ = 0, \text{ almost surely for each } i\in\mathcal{N} \text{ and } j\in\mathcal{M}.$$

Inequality (35) then follows immediately.

- *Step 2.* For problem $\mathcal{OP}_*(\gamma)$, if $\boldsymbol{\alpha}$ is single-period feasible, it holds that

$$\limsup_{\gamma\uparrow+\infty}\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)} y_i^j(t) \le \alpha_i^j \text{ almost surely for all } i\in\mathcal{N} \text{ and } j\in\mathcal{M}. \tag{38}$$

Assume that, to the contrary, there exists (i_0, j_0) such that

$$\limsup_{\gamma\uparrow+\infty}\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)} y_{i_0}^{j_0}(t) > \alpha_{i_0}^{j_0}.$$

Hence, there exists some $\Delta>0$, such that

$$\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)} y_{i_0}^{j_0}(t) > \alpha_{i_0}^{j_0}+\Delta \text{ for infinitely many } \gamma. \tag{39}$$

Denote the set of γ's that satisfy (39) as Γ. Note that $\frac{1}{T(\gamma)}(\sum_{t=1}^{s} y_{i_0}^{j_0}(t))$ increases by at most $1/(T(\gamma))$ as s increases by 1. Hence, for all $\gamma \in \Gamma$ and $\gamma > 3/(T\Delta)$, $\frac{1}{T(\gamma)}(\sum_{t=1}^{s} y_{i_0}^{j_0}(t))$ increases by no more than $\Delta/3$ if s increases by 1. Therefore, for all $\gamma \in \Gamma$ and $\gamma > 3/(T\Delta)$, there exists a $s(\gamma) < T(\gamma)$, such that

$$\alpha_{i_0}^{j_0} + \frac{\Delta}{3} < \frac{1}{T(\gamma)} \sum_{t=1}^{s(\gamma)} y_{i_0}^{j_0}(t) < \alpha_{i_0}^{j_0} + \frac{2\Delta}{3} \tag{40}$$

By (40), we have that, for infinitely many γ,

$$\sum_{t=1}^{s(\gamma)} y_{i_0}^{j_0}(t) > T(\gamma)\left(\alpha_{i_0}^{j_0} + \frac{\Delta}{3}\right).$$

Hence, for infinitely many γ,

$$(d_{i_0}^{j_0}(t))^+ = \left((t-1)\alpha_{i_0}^{j_0} - \sum_{\tau=1}^{t-1} y_{i_0}^{j_0}(\tau)\right)^+ = 0 \text{ for all } t \geq s(\gamma)+1,$$

where the equality follows from

$$\sum_{\tau=1}^{t-1} y_{i_0}^{j_0}(\tau) \geq \sum_{\tau=1}^{s(\gamma)} y_{i_0}^{j_0}(\tau) > T(\gamma)\alpha_{i_0}^{j_0} > (t-1)\alpha_{i_0}^{j_0}.$$

Therefore, ad i_0 will not be offered to customer type j_0 for all $t \geq s(\gamma)+1$. Hence, $y_{i_0}^{j_0}(t) = 0$ for all $t \geq s(\gamma)+1$ and $t \leq T(\gamma)$. By (40), we have

$$\frac{\sum_{t=1}^{T(\gamma)} y_{i_0}^{j_0}(t)}{T(\gamma)} = \frac{\sum_{t=1}^{s(\gamma)} y_{i_0}^{j_0}(t)}{T(\gamma)} < \alpha_{i_0}^{j_0} + \frac{2\Delta}{3} \text{ for } \gamma \in \Gamma \text{ and } \gamma > \frac{3}{T\Delta},$$

which contradicts inequality (39). Therefore, for the system of $\mathcal{OP}_*(\gamma)$, we have inequality (38) holds.

- *Step 3.* For problem $\mathcal{OP}(\gamma)$, if $\boldsymbol{\alpha}$ is first-stage feasible, then (13) holds.

Inequalities (35) and (38) together imply that (13) holds for problem $\mathcal{OP}_*(\gamma)$. In particular, for problem $\mathcal{OP}_*(\gamma)$, we have no stock-out occurs for any ad asymptotically, i.e.,

$$\lim_{\gamma\uparrow+\infty} \frac{1}{T(\gamma)} \sum_{j\in\mathcal{M}} \sum_{t=1}^{T(\gamma)} y_i^j(t) = \sum_{j\in\mathcal{M}} \alpha_i^j \leq \frac{B_i(\gamma)}{b_i T(\gamma)} = \frac{B_i}{b_i T} \text{ for all } i \in \mathcal{N}. \tag{41}$$

Furthermore, by construction, the click-through process of $\mathcal{OP}_*(\gamma)$ is *identical* to that of $\mathcal{OP}(\gamma)$ before stock-out occurs in $\mathcal{OP}(\gamma)$.

We now show that the stock-out probability of any ad's budget converges to 0 for $\mathcal{OP}(\gamma)$ as $\gamma \uparrow +\infty$. If stock-out occurs in $\mathcal{OP}(\gamma)$, there exists some $i \in \mathcal{N}$, such that $b_i \sum_{j\in\mathcal{M}} \sum_{t=1}^{T(\gamma)} y_i^j(t) > B_i(\gamma)$ for $\mathcal{OP}_*(\gamma)$. We have $\frac{1}{T(\gamma)} \sum_{j\in\mathcal{M}} \sum_{t=1}^{T(\gamma)} y_i^j(t) > \frac{B_i(\gamma)}{b_i T(\gamma)} = \frac{B_i}{b_i T} \geq \sum_{j\in\mathcal{M}} \alpha_i^j$ by the feasibility of $\boldsymbol{\alpha}$. Hence, for $\mathcal{OP}_*(\gamma)$,

$$\mathbb{P}\left\{\limsup_{\gamma\uparrow+\infty} \frac{1}{T(\gamma)} \sum_{j\in\mathcal{M}} \sum_{t=1}^{T(\gamma)} y_i^j(t) > \frac{B_i}{b_i T}\right\} \leq \mathbb{P}\left\{\limsup_{\gamma\uparrow+\infty} \frac{1}{T(\gamma)} \sum_{j\in\mathcal{M}} \sum_{t=1}^{T(\gamma)} y_i^j(t) > \sum_{j\in\mathcal{M}} \alpha_i^j\right\} = 0, \tag{42}$$

where the equality follows from (41). Because $\mathcal{OP}_*(\gamma)$ and $\mathcal{OP}(\gamma)$ are equivalent before stock-out occurs, (42) implies that stock-out occurs with probability 0 as $\gamma \uparrow +\infty$ for $\mathcal{OP}(\gamma)$. Therefore, $\mathcal{OP}_*(\gamma)$ and $\mathcal{OP}(\gamma)$ are equivalent with probability 1 as $\gamma \uparrow +\infty$. Since (13) holds for problem $\mathcal{OP}_*(\gamma)$, a standard coupling argument implies that (13) holds for $\mathcal{OP}(\gamma)$ as well. This completes the proof of Theorem 2. □

Before the proof of Theorem 3, we prove Lemma 2 and Proposition 2 first.

Proof of Lemma 2

We first check the feasibility of $\hat{\boldsymbol{\alpha}}(\boldsymbol{z})$ by directly plugging $\hat{\alpha}_i^j(\boldsymbol{z})$ into the constraints of $(\mathcal{OTP})$. Since $\boldsymbol{z}$ is feasible to $(\mathcal{OP}_{\mathsf{Fluid}})$, we have, for all $i \in \mathcal{N}$,

$$b_i \sum_{j\in\mathcal{M}} \hat{\alpha}_i^j(\boldsymbol{z}) = b_i \sum_{j\in\mathcal{M}, S\in\mathfrak{S}^j} p^j \phi_i^j(S) z^j(S) \le \frac{B_i}{T};$$

and, for all $i \in \mathcal{N}$ and $\mathcal{C} \in \mathfrak{K}_i$,

$$\sum_{j\in\mathcal{C}} \hat{\alpha}_i^j(\boldsymbol{z}) = \sum_{j\in\mathcal{C}, S\in\mathfrak{S}^j} p^j \phi_i^j(S) z^j(S) \ge \frac{\eta_i^{\mathcal{C}}}{T}.$$

In addition,

$$\mathcal{V}_{\mathsf{CT}}(\hat{\boldsymbol{\alpha}}(\boldsymbol{z})) = \sum_{i\in\mathcal{N}} \sum_{j\in\mathcal{M}} r_i^j \sum_{S\in\mathfrak{S}^j} \hat{\alpha}_i^j(\boldsymbol{z}) + \lambda F(\hat{\boldsymbol{\alpha}}(\boldsymbol{z})) = \sum_{i\in\mathcal{N}, j\in\mathcal{M}, S\in\mathfrak{S}^j} r_i^j p^j \phi_i^j(S) z^j(S) + \lambda F(\hat{\boldsymbol{\alpha}}(\boldsymbol{z})) = \mathcal{V}_{\mathsf{Fluid}}(\boldsymbol{z}).$$

Therefore, it remains to show that

$$\mathbb{E}[y_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}))] = \hat{\alpha}_i^j(\boldsymbol{z}). \tag{43}$$

Applying the law of total probability, we directly evaluate that

$$\begin{aligned}
\mathbb{E}[y_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}))] =& \mathbb{P}[j(t)=j] \sum_{S\in\mathfrak{S}^j} \mathbb{P}[\pi_{\mathsf{Fluid}}(j)=S|j(t)=j] \mathbb{E}[y_i^j(t)|\pi_{\mathsf{Fluid}}(j)=S, j(t)=j] \\
=& \mathbb{P}[j(t)=j] \sum_{S\in\mathfrak{S}^j} \mathbb{P}[\pi_{\mathsf{Fluid}}(j)=S] \mathbb{E}[y_i^j(t)|\pi_{\mathsf{Fluid}}(j)=S, j(t)=j] \\
=& p^j \sum_{S\in\mathfrak{S}^j} z^j(S)\phi_i^j(S) \\
=& \hat{\alpha}_i^j(\boldsymbol{z}),
\end{aligned}$$

i.e., (43) holds. Therefore, the click-through target vector $\hat{\boldsymbol{\alpha}}(\boldsymbol{z})$ is first-stage feasible. In particular, $\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)$ is first-stage feasible with $\mathbb{E}[y_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*))] = \hat{\alpha}_i^j(\boldsymbol{z}^*)$. We defer the proof of $\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)$'s optimality for $(\mathcal{OTP})$ to the proof of Theorem 3. □

Before proving Theorem 3, we present the proof of Proposition 2 first.

Proof of Proposition 2

We prove (17) by showing each individual equality or inequality thereof.

- *Step 1.* The FV generated by $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)$ in $\mathcal{OP}(\gamma)$ is asymptotically identical to the optimal FV in $(\mathcal{OP}_{\mathsf{Fluid}})$, i.e.,

$$\lim_{\gamma\uparrow+\infty} \mathcal{V}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)|\gamma) = \mathcal{V}^*_{\mathsf{Fluid}}. \tag{44}$$

We first show that $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)$ is asymptotically *feasible* for $\mathcal{OP}(\gamma)$. As $\gamma \uparrow +\infty$, we have

$$\lim_{\gamma\uparrow+\infty} \frac{1}{T(\gamma)} \sum_{t=1}^{T(\gamma)} \sum_{j\in\mathcal{M}} b_i y_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)) = \mathbb{E}\left[\sum_{j\in\mathcal{M}} b_i y_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*))\right] = \sum_{j\in\mathcal{M}, S\in\mathfrak{S}^j} b_i p^j \phi_i^j(S) z^{j*}(S) \le \frac{B_i(\gamma)}{T(\gamma)}, \tag{45}$$

where the first equality follows from the strong law of large numbers, the second from the definition of $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)$, and the inequality follows from $\boldsymbol{z}^*$ is feasible for $(\mathcal{OP}_{\mathsf{Fluid}})$. Similarly, we have, under the policy $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)$,

$$\lim_{\gamma\uparrow+\infty} \mathbb{E}\left[\frac{1}{T(\gamma)} \sum_{t=1}^{T(\gamma)} \sum_{j\in\mathcal{C}} y_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*))\right] = \sum_{j\in\mathcal{C}, S\in\mathfrak{S}^j} p^j \phi_i^j(S) z^j(S) \ge \frac{\eta_i^{\mathcal{C}}(\gamma)}{T(\gamma)}, \tag{46}$$

where the equality follows from the definition of $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)$, and the inequality follows from $\boldsymbol{z}^*$ is feasible for $(\mathcal{OP}_{\mathsf{Fluid}})$. Inequalities (45) and (46) together imply that $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)$ is asymptotically feasible for $\mathcal{OP}(\gamma)$ as $\gamma \uparrow +\infty$.

Then, we evaluate the FV of the Fluid-$\boldsymbol{z}^*$ policy as follows:

$$\begin{aligned}\lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)|\gamma) &= \lim_{\gamma\uparrow+\infty}\mathbb{E}\left[\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^j y_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)) + \lambda F(\bar{\boldsymbol{y}}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)))\right]\\ &= \sum_{i\in\mathcal{N},j\in\mathcal{M},S\in\mathfrak{S}^j} r_i^j p^j \phi_i^j(S) z^{j*}(S) + \lambda F(\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*))\\ &= \mathcal{V}_{\mathsf{Fluid}}(\boldsymbol{z}^*)\\ &= \mathcal{V}^*_{\mathsf{Fluid}},\end{aligned} \tag{47}$$

where the second equality follows from the law of large numbers and the dominated convergence theorem. Therefore, (44) follows from (47).

- *Step 2.* The following inequality holds:

$$\lim_{\gamma\uparrow+\infty}\mathcal{V}^*(\gamma) \geq \lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)|\gamma). \tag{48}$$

Since $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)$ is asymptotically feasible, the optimality of $\mathcal{V}^*(\gamma)$ implies that $\lim_{\gamma\uparrow+\infty}\mathcal{V}^*(\gamma) \geq \lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)|\gamma)$.

- *Step 3.* The optimal FV of $(\mathcal{OP}_{\mathsf{Fluid}})$ dominates that of $\mathcal{OP}(\gamma)$, i.e.,

$$\mathcal{V}^*_{\mathsf{Fluid}} \geq \mathcal{V}^*(\gamma) \text{ for any } \gamma > 0. \tag{49}$$

Consider an arbitrary policy $\pi \in \Pi$ feasible for $\mathcal{OP}(\gamma)$. We first define the following probability measure induced by π, $\boldsymbol{z}_{\mathsf{Fluid}}(\pi)$ for $(\mathcal{OP}_{\mathsf{Fluid}})$: For $j \in \mathcal{M}$ and $S \in \mathfrak{S}^j$,

$$z^j_{\mathsf{Fluid}}(S|\pi) := \mathbb{P}\left[\pi_{\tilde{t}}(j(\tilde{t}), \mathcal{H}_{\tilde{t}-1}) = S | j(\tilde{t}) = j\right], \tag{50}$$

where $\tilde{t}$ is a random variable uniformly distributed on $\{1, 2, ..., T(\gamma)\}$ and independent of everything else. Because π is feasible for $\mathcal{OP}(\gamma)$, all the constraints of $\mathcal{OP}(\gamma)$ will also be satisfied in the expected sense as well, i.e.,

$$\frac{1}{T(\gamma)}\mathbb{E}\left[\sum_{t=1}^{T(\gamma)}\sum_{j\in\mathcal{M}} b_i y_i^j(t|\pi)\right] \leq \frac{B_i(\gamma)}{T(\gamma)}, \text{ for each } i \in \mathcal{N},$$

and

$$\frac{1}{T(\gamma)}\mathbb{E}\left[\sum_{t=1}^{T}\sum_{j\in\mathcal{C}} y_i^j(t|\pi)\right] \geq \frac{\eta_i^{\mathcal{C}}(\gamma)}{T(\gamma)}, \text{ for each } i \in \mathcal{N} \text{ and } \mathcal{C} \in \mathfrak{K}_i,$$

where the expectations are taken with respect to $j(t)$, π, and $\boldsymbol{y}$. By (50), straightforward algebraic manipulation and the law of iterated expectations together yield that

$$\mathbb{E}[\bar{y}_i^j] = \mathbb{E}\left[\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)} y_i^j(t|\pi)\right] = \sum_{S\in\mathfrak{S}^j} p^j \phi_i^j(S) z^j_{\mathsf{Fluid}}(S|\pi) = \hat{\alpha}_i^j(\boldsymbol{z}_{\mathsf{Fluid}}(S|\pi)).$$

Plugging this identity into the constraints of $(\mathcal{OP}_{\mathsf{Fluid}})$, we have

$$\sum_{j\in\mathcal{M},S\in\mathfrak{S}^j} b_i p^j \phi_i^j(S) z^j_{\mathsf{Fluid}}(S|\pi) \leq \frac{B_i(\gamma)}{T(\gamma)} \text{ for each } i \in \mathcal{N}$$

and that

$$\sum_{j\in\mathcal{C},S\in\mathfrak{S}^j} p^j\phi_i^j(S)z^j_{\mathsf{Fluid}}(S|\pi)\geq \frac{\eta_i^{\mathcal{C}}(\gamma)}{T(\gamma)} \text{ for each } i\in\mathcal{N} \text{ and } \mathcal{C}\in\mathfrak{K}_i.$$

Therefore, $\boldsymbol{z}_{\mathsf{Fluid}}(\pi)$ is feasible for $(\mathcal{OP}_{\mathsf{Fluid}})$ and, hence,

$$\mathcal{V}^*_{\mathsf{Fluid}}\geq \mathcal{V}_{\mathsf{Fluid}}(\boldsymbol{z}_{\mathsf{Fluid}}(\pi)). \tag{51}$$

By Jensen's inequality and the concavity of the fairness metric $F(\cdot)$,

$$\begin{aligned}\mathcal{V}_{\mathsf{Fluid}}(\boldsymbol{z}_{\mathsf{Fluid}}(\pi)) &= \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}}\sum_{S\in\mathfrak{S}^j} r_i^j p^j\phi_i^j(S)z^j_{\mathsf{Fluid}}(S|\pi)+\lambda F(\hat{\boldsymbol{\alpha}}(\boldsymbol{z}_{\mathsf{Fluid}}(S|\pi)))\\ &\geq \mathbb{E}\left[\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^j y_i^j(t|\pi)+\lambda F(\bar{\boldsymbol{y}}(\pi))\right] \\ &= \mathcal{V}(\pi|\gamma)\end{aligned}\tag{52}$$

Since π is arbitrary, inequalities (51) and (52) together imply that (49) holds.

Therefore, putting the (in)equalities (44), (48), and (49) together, we have (17) holds, which completes the proof of Proposition 2. □

We are now ready to prove Theorem 3.

Proof of Theorem 3

We prove (14) by the following three steps.

- *Step 1.* The following inequality holds:

$$\mathcal{V}^*_{\mathsf{CT}}\geq \lim_{\gamma\uparrow+\infty}\mathcal{V}^*(\gamma). \tag{53}$$

By Proposition 2, there exists one optimal static policy $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)$ of $(\mathcal{OP})$ in the asymptotic regime. By Lemma 2, we have

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)=(\hat{\alpha}_i^j(\boldsymbol{z}^*), i\in\mathcal{N}, j\in\mathcal{M})\in\mathbb{R}_+^{(nm)}.$$

Therefore, there exists a static policy $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)$, such that

$$\begin{aligned}&\mathbb{E}[y_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*))]\geq\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*), \text{ for each } i\in\mathcal{N}, j\in\mathcal{M},\\ &b_i\sum_{j\in\mathcal{M}}\hat{\alpha}_i^j(\boldsymbol{z}^*)\leq\frac{B_i}{T}, \text{ for each } i\in\mathcal{N},\\ &\sum_{j\in\mathcal{C}}\hat{\alpha}_i^j(\boldsymbol{z}^*)\geq\frac{\eta_i^{\mathcal{C}}}{T}, \text{ for each } i\in\mathcal{N} \text{ and } \mathcal{C}\in\mathfrak{K}_i,\end{aligned}$$

hence $\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)$ is a feasible solution of the first-stage click-through target optimization problem $(\mathcal{OTP})$. By the law of large numbers and the dominated convergence theorem, we obtain $\mathcal{V}^*_{\mathsf{CT}}\geq\lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)|\gamma)$ as follows:

$$\begin{aligned}\mathcal{V}^*_{\mathsf{CT}}&\geq\mathcal{V}_{\mathsf{CT}}(\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*))\\ &=\sum_{i\in\mathcal{N},j\in\mathcal{M}} r_i^j\hat{\alpha}_i^j(\boldsymbol{z}^*)+\lambda F(\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*))\\ &=\lim_{\gamma\uparrow+\infty}\mathbb{E}\left[\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^j y_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*))\right]+\lambda F\left(\mathbb{E}\left[\bar{\boldsymbol{y}}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*))\right]\right)\\ &=\lim_{\gamma\uparrow+\infty}\mathbb{E}\left[\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^j y_i^j(t|\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*))+\lambda F(\bar{\boldsymbol{y}}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)))\right]\\ &=\lim_{\gamma\uparrow+\infty}\mathcal{V}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)|\gamma).\end{aligned}\tag{54}$$

Since $\lim_{\gamma\uparrow+\infty} \mathcal{V}(\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*)|\gamma) = \lim_{\gamma\uparrow+\infty} \mathcal{V}^*(\gamma)$ by Proposition 2, (53) follows.

- *Step 2.* The following inequality holds:

$$\lim_{\gamma\uparrow+\infty} \mathcal{V}^*(\gamma) \geq \lim_{\gamma\uparrow+\infty} \mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma). \tag{55}$$

The optimality of $\mathcal{V}^*(\gamma)$ implies that $\mathcal{V}^*(\gamma) \geq \mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma)$ for all $\gamma > 0$.

- *Step 3.* The DWO-$\boldsymbol{\alpha}^*$ policy generates the same asymptotic FV in $(\mathcal{OP})$ as the optimal FV in $(\mathcal{OTP})$, i.e.,

$$\lim_{\gamma\uparrow+\infty} \mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma) = \mathcal{V}^*_{\mathsf{CT}} \tag{56}$$

By Theorem 2, the DWO-$\boldsymbol{\alpha}^*$ policy is asymptotically feasible for the original problem $(\mathcal{OP})$ with (13) holding true. We now evaluate $\mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma)$ in the asymptotic regime:

$$\begin{aligned}
\lim_{\gamma\uparrow+\infty} \mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma) &= \lim_{\gamma\uparrow+\infty} \mathbb{E}\left[\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^j y_i^j(t|\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)) + \lambda F(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))\right]\\
&= \sum_{i\in\mathcal{N},j\in\mathcal{M}} r_i^j \alpha_i^{j*} + \lambda F(\boldsymbol{\alpha}^*)\\
&= \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}^*)\\
&= \mathcal{V}^*_{\mathsf{CT}},
\end{aligned} \tag{57}$$

where the second equality follows from equality (13) and the dominated convergence theorem. Hence, equality (56) follows from equality (57). Therefore, putting the (in)equalities (53), (55), and (56) together, we have (14) holds. As a by-product of the proof, (14) also implies

$$\mathcal{V}^*_{\mathsf{CT}} = \lim_{\gamma\uparrow+\infty} \mathcal{V}^*(\gamma) = \mathcal{V}^*_{\mathsf{Fluid}} = \mathcal{V}_{\mathsf{CT}}(\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)).$$

By (49) in the proof of Proposition 2, we have

$$\mathcal{V}^*_{\mathsf{CT}} = \mathcal{V}^*_{\mathsf{Fluid}} \geq \mathcal{V}^*(\gamma) \text{ for any } \gamma > 0.$$

Hence, we have (15) holds. Moreover, $\hat{\boldsymbol{\alpha}}(\boldsymbol{z}^*)$ is optimal for $(\mathcal{OTP})$, which completes the proof of Lemma 2.

To complete the proof, we now show (16). Because (15) holds for any $\gamma > 0$, we have

$$\begin{aligned}
&\mathcal{V}^*(\gamma) - \mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma) \leq \mathcal{V}^*_{\mathsf{CT}} - \mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma)\\
&= \left[\sum_{i\in\mathcal{N},j\in\mathcal{M}} r_i^j\alpha_i^{j*} + \lambda F(\boldsymbol{\alpha}^*)\right] - \mathbb{E}\left[\frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{M}} r_i^j y_i^j(t|\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)) + \lambda F(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))\right]\\
&= \sum_{i\in\mathcal{N},j\in\mathcal{M}} r_i^j \mathbb{E}\left[\alpha_i^{j*} - \frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)} y_i^j(t|\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*))\right] + \lambda\left(F(\boldsymbol{\alpha}^*) - \mathbb{E}[F(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))]\right)
\end{aligned} \tag{58}$$

By the definition of the debt vector $\boldsymbol{d}(t)$, we can bound the first term of (58) as follows

$$\begin{aligned}
\sum_{i\in\mathcal{N},j\in\mathcal{M}} r_i^j \mathbb{E}\left[\alpha_i^{j*} - \frac{1}{T(\gamma)}\sum_{t=1}^{T(\gamma)} y_i^j(t|\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*))\right] &= \frac{1}{T(\gamma)}\sum_{i\in\mathcal{N},j\in\mathcal{M}} r_i^j \mathbb{E}\left[d_i^j(T(\gamma)+1)\right]\\
&\leq \frac{1}{T(\gamma)}\sum_{i\in\mathcal{N},j\in\mathcal{M}} r_i^j \mathbb{E}\left[(d_i^j(T(\gamma)+1))^+\right]\\
&\leq \frac{1}{T(\gamma)}||\boldsymbol{r}||_2 \cdot ||\mathbb{E}\left[(\boldsymbol{d}(T(\gamma)+1))^+\right]||_2\\
&\leq \frac{\mathcal{C}_1}{\sqrt{\gamma}}
\end{aligned} \tag{59}$$

where the constant $\mathcal{C}_1 := ||\boldsymbol{r}||_2 \cdot \sqrt{\frac{2mn}{T}}$, the second inequality follows from the Cauchy–Schwarz inequality, and the last from (37).

Since $F(\cdot)$ is a concave function, it has subgradient $\boldsymbol{f}(\cdot) := (f_i^j(\cdot) : i \in \mathcal{N}, j \in \mathcal{M}) \in \mathbb{R}^{nm}$. Define $f_{\mathsf{max}} := \max_{i\in\mathcal{N}, j\in\mathcal{M}, \boldsymbol{y}\in[0,1]^{nm}} |f_i^j(\boldsymbol{y})|$ and $F_{\mathsf{max}} := \max_{\boldsymbol{\alpha}\in\in[0,1]^{nm}} |F(\boldsymbol{\alpha})|$. We have

$$F(\boldsymbol{\alpha}^*) \le F(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*))) + \boldsymbol{f}(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))^{\intercal}(\boldsymbol{\alpha}^* - \bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*))).$$

Hence, we can bound the second term of (58) as follows, for $\gamma \ge 1$,

$$\begin{aligned} F(\boldsymbol{\alpha}^*) - \mathbb{E}[F(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))] &\le \mathbb{E}[\boldsymbol{f}(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))^{\intercal}(\boldsymbol{\alpha}^* - \bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))] \\ &\le \frac{1}{T(\gamma)}\sqrt{\mathbb{E}||\boldsymbol{f}(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))||_2^2} \cdot \sqrt{\mathbb{E}||\boldsymbol{d}(T(\gamma)+1)||_2^2} \\ &\le \frac{1}{T(\gamma)}\sqrt{\mathbb{E}||\boldsymbol{f}(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))||_2^2} \cdot \sqrt{\mathbb{E}||(\boldsymbol{d}(T(\gamma)+1))^+||_2^2 + mn} \\ &\le \frac{\sqrt{nm}\cdot f_{\mathsf{max}}}{T(\gamma)} \cdot \sqrt{2nmT(\gamma)+mn} \\ &\le \frac{\mathcal{C}_2}{\sqrt{\gamma}}, \end{aligned} \tag{60}$$

where the constant $\mathcal{C}_2 := \max\{\frac{\sqrt{nm}\cdot f_{\mathsf{max}}}{T} \cdot \sqrt{2nmT+mn}, F_{\mathsf{max}}\}$, the second inequality follows from the Cauchy–Schwarz inequality, the third from the fact that

$$d_i^j(T(\gamma)+1))^- = \left(T(\gamma)\alpha_i^{j*} - \sum_{t=1}^{T(\gamma)} y_i^j(t|\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*))\right)^- \le 1, \tag{61}$$

and the fourth from (36). Inequality (61) holds because, under the DWO-$\boldsymbol{\alpha}^*$ policy, once $\sum_{t=1}^{s} y_i^j(t|\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)) \ge T(\gamma)\alpha_i^{j*}$ for some $s \le T(\gamma)$, ad i will not be offered to type j customer for all $t \ge s+1$. For $\gamma < 1$, it holds that

$$F(\boldsymbol{\alpha}^*) - \mathbb{E}[F(\bar{\boldsymbol{y}}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)))] \le \frac{F_{\mathsf{max}}}{\sqrt{\gamma}} \le \frac{\mathcal{C}_2}{\sqrt{\gamma}}. \tag{62}$$

Combining (59), (60), and (62) yields that

$$\mathcal{V}^*(\gamma) - \mathcal{V}(\pi_{\mathsf{DWO}}(\boldsymbol{\alpha}^*)|\gamma) \le \frac{\mathcal{C}_1 + \lambda\cdot\mathcal{C}_2}{\sqrt{\gamma}}.$$

Hence, we have (16) holds with $\mathcal{C} := \mathcal{C}_1 + \lambda\cdot\mathcal{C}_2 > 0$. This completes the proof of Theorem 3. □

Proof of Proposition 3

The proof follows from the same argument as *Step 3* in the proof of Theorem 3 by replacing $\boldsymbol{\alpha}^*$ with any feasible $\boldsymbol{\alpha}$. To avoid repetition, we omit the proof details. □

E. Feasible Click-Through Targets Under the MNL Choice Model

Proposition 1 characterizes the feasible region of the click-through targets $\mathcal{A}_{MNL}$ if customers follow the MNL choice model. This section seeks to deliver additional insights on when the click-through targets are feasible. We observe that (34) is equivalent to

$$p^j \ge \sum_{i\in\mathcal{N}} \alpha_i^j + \max\left\{\frac{1}{K}\sum_{i\in\mathcal{N}}\frac{\alpha_i^j}{v_i^j}, \max_{i\in\mathcal{N}}\left\{\frac{\alpha_i^j}{v_i^j}\right\}\right\}, \text{ for each } j \in \mathcal{M}.$$

Here, p^j is the expected (per-user) traffic of type-j customers in each period. Clearly, $\sum_{i\in\mathcal{N}} \alpha_i^j$ is the total required traffic for type-j customers if a customer will click one of the ad in the offer set with probability 1. In practice, however, a customer may end up not choosing any ad from the offer set, so we need some buffer traffic for type-j customers that accounts for the non-click circumstance.

More specifically, let $\mathfrak{S}_i$ denote the collection of all offer sets containing ad i. Since the offer-set policy may be random, we define $\mu_j(S)$ as the probability of displaying offer-set $S\subseteq\mathcal{N}$ to type-j customers. Thus, the desired click-through goal for ad i and type-j customer is

$$\sum_{S\in\mathfrak{S}_i} \mu_j(S)\cdot \frac{v_i^j}{1+\sum_{i'\in S} v_{i'}^j} \geq \alpha_i^j.$$

Thus, the non-click probability of the ads for a type-j customer when ad i ($i\in\mathcal{N}$) is offered satisfies that

$$\alpha_i^j(o) := \sum_{S\in\mathfrak{S}_i} \mu_j(S)\cdot \frac{1}{1+\sum_{i'\in S} v_{i'}^j} \geq \frac{\alpha_i^j}{v_i^j}.$$

Therefore, to ensure the click-through goal of type-j customers and ad i, the traffic of customer type j must satisfy $p^j \geq \sum_{i'\in\mathcal{N}} \alpha_{i'}^j + \alpha_i^j(o) \geq \sum_{i'\in\mathcal{N}} \alpha_{i'}^j + \frac{\alpha_i^j}{v_i^j}$ for all $i\in\mathcal{N}$.

The cardinality constraint for the offer set size would impose an additional bound on the non-click probability of type-j customers. Specifically, let $\mathfrak{S} := \bigcup_{i=1}^n \mathfrak{S}_i$ be the set of all offer sets displayed to a customer. Because $|S|\leq K$ for any $S\in\mathfrak{S}$, $|\{i\in\mathcal{N} : S\in\mathfrak{S}_i\}|\leq K$ for all S. We have, given customer type j,

$$K\sum_{S\in\mathfrak{S}} \mu_j(S)\cdot\frac{1}{1+\sum_{i\in S} v_i^j} \geq \sum_{i\in\mathcal{N}}\sum_{S\in\mathfrak{S}_i} \mu_j(S)\cdot \frac{1}{1+\sum_{i\in S} v_i^j} \geq \sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{v_i^j}.$$

Thus, the non-click probability of all ads for type-j customer satisfies that

$$\alpha_o^j := \sum_{S\in\mathfrak{S}} \mu_j(S)\cdot \frac{1}{1+\sum_{i\in S} v_i^j} \geq \frac{1}{K}\sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{v_i^j}.$$

Therefore, given the cardinality constraint of an offer set, to ensure the click-through targets of type-j customers with respect to all ads, the traffic of customer type j must satisfy $p^j \geq \sum_{i\in\mathcal{N}} \alpha_i^j + \alpha_o^j \geq \sum_{i\in\mathcal{N}} \alpha_i^j + \frac{1}{K}\sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{v_i^j}$. In summary, the characterization for the feasibility of $\boldsymbol{\alpha}$ demonstrates that, to meet the click-through targets, we should also account for the *non-click* cases.

F. Optimal Target Convex Program Formulation for Specific Choice Models

In this section, we introduce the characterization of first-stage feasible region $\mathcal{A} := \{\boldsymbol{\alpha}\in[0,1]^{nm} : h(\boldsymbol{\alpha})\geq 0\}$ for independent and generalized attraction choice models. Similar to the characterization of $\mathcal{A}$ if the customers follow the MNL model, we use a binary variable representation of a deterministic static offer-set policy $\pi\in\Pi_{\mathsf{d}}$ (see, also, the proof of Proposition 1). More specifically, $\pi\in\Pi_{\mathsf{d}}$ can be equivalently represented by an $nm-$dimensional binary vector $\boldsymbol{x} = (x_i^j\in\{0,1\} : i\in\mathcal{N}, j\in\mathcal{M})$, where $x_i^j = 1$ means that $i\in\pi(j)$, i.e., ad i is included in the offer-set displayed to a type-j customer. We denote $\phi_i^j(\boldsymbol{x})$ as the expected click-throughs of a type-j customer for ad i if the offer-set displayed to this customer is $S^j = \{i\in\mathcal{N} : x_i^j = 1\}$. Denote the set of all plausible offer-set representation vectors as $\mathcal{X}\subset\{0,1\}^{nm}$, and the set of plausible offer-set representation vectors displayed to a type-j customer as $\mathcal{X}^j\subset\{0,1\}^n$.

F.1. Independent Choice Model

If customers follow the independent choice model, the click-throughs only depend on the customer type j and ad i, but not on the offer-set S^j displayed to the customer. This is actually the most commonly adopted choice models in practical advertising (see, e.g., Feldman et al. 2022), where the click-through rate (CTR) prediction algorithm of the platform outputs:

$$\phi_i^j(\boldsymbol{x}) = c_i^j x_i^j, \tag{63}$$

where $c_i^j > 0$ is the CTR of ad i with respect to a type-j customer. It is evident from (63) that the number of click-throughs is independent of the ads in S^j other than ad i. The following proposition is a counterpart of Proposition 1 with the independent choice model.

Proposition 4. *If customers follow the independent click-through model* (63)*, the first-stage feasible region $\mathcal{A}_{IND}$ is given by the following linear constraints:*

$$\mathcal{A}_{IND} = \left\{\boldsymbol{\alpha} \in [0,1]^{nm} : p^j c_i^j \geq \alpha_i^j, \ \textit{for each } i \in \mathcal{N}, j \in \mathcal{M}\right\}. \tag{64}$$

It is expected that, under the independent choice model, we have an independent feasibility condition for each ad-customer type pair. With Proposition 4, we we can simplify the optimal target problem $(\mathcal{OTP})$ under the independent choice model as the following convex program with linear constraints only:

$$\begin{aligned} \max_{\boldsymbol{\alpha} \geq \mathbf{0}} \ & \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) \\ \text{s.t. } & p^j c_i^j \geq \alpha_i^j, \ \text{for each } i \in \mathcal{N}, j \in \mathcal{M} \\ & b_i \sum_{j \in \mathcal{M}} \alpha_i^j \leq \frac{B_i}{T}, \ \text{for each } i \in \mathcal{N}, \\ & \sum_{j \in \mathcal{C}} \alpha_i^j \geq \frac{\eta_i^{\mathcal{C}}}{T}, \ \text{for each } i \in \mathcal{N} \text{ and } \mathcal{C} \in \mathfrak{K}_i. \end{aligned} \tag{$\mathcal{OTP}-\mathcal{IND}$}$$

F.2. Generalized Attraction Model

The generalized attraction model (GAM) is a generalization of MNL accounts for the possibility that a customer may look for a product outside the offer-set (see, e.g., Gallego et al. 2015). Under the GAM, the expected number of click-throughs of ad i by a type-j customer is given by

$$\phi_i^j(\boldsymbol{x}) = \frac{v_i^j x_i^j}{1 + \sum_{i' \in \mathcal{N}} \omega_{i'}^j (1 - x_{i'}^j) + \sum_{i \in \mathcal{N}} v_{i'}^j x_{i'}^j}, \tag{65}$$

where $v_i^j > 0$ is the attraction value of ad i to a type-j customer, and $w_i^j \in [0, v_i^j]$ is the shadow attraction value of ad i to a type-j customer, capturing the customer's looking for a product outside the offer-set. Hence, by defining $\tilde{v}^j := 1 + \sum_{i \in \mathcal{N}} \omega_i^j > 0$ and $\tilde{v}_i^j := v_i^j - \omega_i^j \geq 0$, we have, under the GAM,

$$\phi_i^j(\boldsymbol{x}) = \frac{v_i^j x_i^j}{\tilde{v}^j + \sum_{i' \in \mathcal{N}} \tilde{v}_{i'}^j x_{i'}^j}. \tag{66}$$

As Corollary 1, we first rewrite the necessary and sufficient condition for $\boldsymbol{\alpha}$ under the GAM.

COROLLARY 2. *If customers follow the GAM click-through model (66), a click-through target vector $\boldsymbol{\alpha}$ is single-period feasible if and only if, for each $j \in \mathcal{M}$*

$$\max_{\boldsymbol{x}^j \in \mathcal{X}^j} \sum_{i \in \mathcal{N}} \frac{p^j v_i^j \theta_i^j x_i^j}{\tilde{v}^j + \sum_{i' \in \mathcal{N}} \tilde{v}_{i'}^j x_{i'}^j} \geq \sum_{i \in \mathcal{N}} \alpha_i^j \theta_i^j \text{ for any } \theta_i^j \geq 0 \ (i \in \mathcal{N}). \tag{67}$$

Similar to Proposition 1, we can characterize the first-stage feasible region of the click-through target vector under the GAM, $\mathcal{A}_{GAM} := \{\boldsymbol{\alpha} : (67) \text{ holds for each } j \in \mathcal{M}\}$, using linear constraints. The following proposition characterizes $\mathcal{A}_{GAM}$ and accounts for the offer-set cardinality constraint.

PROPOSITION 5. *If customers follow the GAM (66) and the set of all feasible offer-sets is $\mathfrak{S}^j = \{S \subset \mathcal{N} : |S| \leq K\}$ for each $j \in \mathcal{M}$, the first-stage feasible region $\mathcal{A}_{GAM}$ is given by the following linear constraints:*

$$\mathcal{A}_{GAM} := \left\{ \boldsymbol{\alpha} \in \mathbb{R}_+^{nm} : \sum_{i' \in \mathcal{N}} \frac{\tilde{v}_{i'}^j \alpha_{i'}^j}{v_{i'}^j} + \frac{\tilde{v}^j \alpha_i^j}{v_i^j} \leq p^j, \text{ for each } i \in \mathcal{N}, j \in \mathcal{M}, \text{ and } \sum_{i \in \mathcal{N}} \frac{\tilde{v}_i^j \alpha_i^j}{v_i^j} + \frac{\tilde{v}^j}{K} \sum_{i \in \mathcal{N}} \frac{\alpha_i^j}{v_i^j} \leq p^j, \text{ for each } j \in \mathcal{M} \right\}. \tag{68}$$

With Proposition 5, we can simplify the optimal target problem $(\mathcal{OTP})$ under the generalized attraction model as the following convex program with linear constraints only:

$$\begin{aligned}
\max_{\boldsymbol{\alpha} \geq \mathbf{0}} \ & \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) \\
\text{s.t. } & \sum_{i' \in \mathcal{N}} \frac{\tilde{v}_{i'}^j \alpha_{i'}^j}{v_{i'}^j} + \frac{\tilde{v}^j \alpha_i^j}{v_i^j} \leq p^j, \text{ for each } i \in \mathcal{N}, j \in \mathcal{M}, \\
& \sum_{i \in \mathcal{N}} \frac{\tilde{v}_i^j \alpha_i^j}{v_i^j} + \frac{\tilde{v}^j}{K} \sum_{i \in \mathcal{N}} \frac{\alpha_i^j}{v_i^j} \leq p^j, \text{ for each } j \\
& b_i \sum_{j \in \mathcal{M}} \alpha_i^j \leq \frac{B_i}{T}, \text{ for each } i \in \mathcal{N}, \\
& \sum_{j \in \mathcal{C}} \alpha_i^j \geq \frac{\eta_i^{\mathcal{C}}}{T}, \text{ for each } i \in \mathcal{N} \text{ and } \mathcal{C} \in \mathfrak{K}_i.
\end{aligned} \tag{$\mathcal{OTP}-\mathcal{GAM}$}$$

F.3. Proofs

In this subsection, we give proofs of the technical results presented in Appendix F.

Proof of Proposition 4

Directly applying Theorem 1 to the independent choice model implies that $\boldsymbol{\alpha}$ is feasible if and only if the following inequality holds.

$$\max_{\boldsymbol{x} \in \mathcal{X}} \sum_{i \in \mathcal{N}, j \in \mathcal{M}} p^j c_i^j \theta_i^j x_i^j \geq \sum_{i \in \mathcal{N}, j \in \mathcal{M}} \alpha_i^j \theta_i^j \text{ for any } \boldsymbol{\theta} \geq \mathbf{0}. \tag{69}$$

Setting $\theta_i^j = 1$ and all other θ's equal to zero in (69) immediately implies that if $\boldsymbol{\alpha}$ is single-period feasible, then (64) holds. Reversely, if (64) holds, then for any $\boldsymbol{\theta} \geq \mathbf{0}$, we have

$$\max_{\boldsymbol{x} \in \mathcal{X}} \sum_{i \in \mathcal{N}, j \in \mathcal{M}} p^j c_i^j \theta_i^j x_i^j = \sum_{i \in \mathcal{N}, j \in \mathcal{M}} p^j c_i^j \theta_i^j \geq \sum_{i \in \mathcal{N}, j \in \mathcal{M}} \alpha_i^j \theta_i^j,$$

i.e., (69) holds and, therefore, $\boldsymbol{\alpha}$ is single-period feasible. This concludes the proof of Proposition 4. $\square$

Proof of Corollary 2

The proof follows from exactly the same argument as that of Corollary 1. We omit the details to avoid repetition. □

Proof of Proposition 5

As in the proof of Proposition 1, we first state and prove the following auxiliary lemma.

LEMMA 4. *If customers follow the GAM click-through model (66) and the set of all feasible offer-sets is $\mathfrak{S}^j = \{S \subset \mathcal{N} : |S| \le K\}$ for each $j \in \mathcal{M}$, we have $\boldsymbol{\alpha}$ is single-period feasible if and only if there exist $\boldsymbol{w} := (w_i^j : i \in \mathcal{N}, j \in \mathcal{M})$ and $\boldsymbol{z} := (z^j : j \in \mathcal{M})$ that satisfy the following linear constraints*

$$\begin{aligned} &p^j v_i^j w_i^j \ge \alpha_i^j,\ w_i^j \le z^j,\ w_i^j \ge 0,\ \textit{for each } i \in \mathcal{N}, j \in \mathcal{M}, \\ &\sum_{i\in\mathcal{N}} \tilde{v}_i^j w_i^j + \tilde{v}^j z^j = 1,\ \sum_{i\in\mathcal{N}} w_i^j \le K z^j,\ \textit{for each } j \in \mathcal{M}, \end{aligned} \tag{70}$$

where $z^j := \frac{1}{\tilde{v}^j + \sum_{i'\in\mathcal{N}} \tilde{v}_{i'}^j x_{i'}^j}$ and $w_i^j := x_i^j z^j = \frac{x_i^j}{\tilde{v}^j + \sum_{i'\in\mathcal{N}} \tilde{v}_{i'}^j x_{i'}^j}$.

Proof of Lemma 4 Similar to the proof of Lemma 3, it is straightforward to check that the left-hand side of (67) is quasi-convex in $\boldsymbol{x}^j$ for all j, so there always exists a maximizer on the boundary of the feasible region. Thus, we can relax the binary constraint $x_i^j \in \{0,1\}$ to $x_i^j \in [0,1]$ in (67), which is therefore equivalent to

$$\max_{\boldsymbol{x}^j \in [0,1]^n,\ \sum_{i\in\mathcal{N}} x_i^j \le K} \sum_{i\in\mathcal{N}} \frac{p^j v_i^j \theta_i^j x_i^j}{\tilde{v}^j + \sum_{i'\in\mathcal{N}} \tilde{v}_{i'}^j x_{i'}^j} \ge \sum_{i\in\mathcal{N}} \alpha_i^j \theta_i^j \text{ for all } \boldsymbol{\theta}^j \ge \mathbf{0} \text{ and } j \in \mathcal{M}. \tag{71}$$

We change the decision variable and define, for all $j \in \mathcal{M}$,

$$z^j := \frac{1}{\tilde{v}^j + \sum_{i'\in\mathcal{N}} \tilde{v}_{i'}^j x_{i'}^j} \text{ and } w_i^j := x_i^j z^j = \frac{x_i^j}{\tilde{v}^j + \sum_{i'\in\mathcal{N}} \tilde{v}_{i'}^j x_{i'}^j}.$$

Then, we can rewrite (71) as, for any j,

$$\begin{aligned} \min_{\boldsymbol{\theta}^j \ge \mathbf{0}} &\left(\max_{w_i^j, z^j} \sum_{i\in\mathcal{N}} p^j v_i^j w_i^j \theta_i^j - \sum_{i\in\mathcal{N}} \alpha_i^j \theta_i^j \right) \ge 0 \\ \text{s.t. } &\sum_{i\in\mathcal{N}} \tilde{v}_i^j w_i^j + \tilde{v}^j z^j = 1, \\ &\sum_{i\in\mathcal{N}} w_i^j \le K z^j, \\ &0 \le w_i^j \le z^j \text{ for each } i \in \mathcal{N}. \end{aligned} \tag{72}$$

By Sion's minimax theorem, we can exchange the maximization and minimization operators so that (72) is equivalent to, for any $j \in \mathcal{M}$:

$$\begin{aligned} \max_{\boldsymbol{w}^j, z^j} \min_{\boldsymbol{\theta}^j \ge 0} &\sum_{i\in\mathcal{N}} \theta_i^j (p^j v_i^j w_i^j - \alpha_i^j) \ge 0, \\ \text{s.t. } &\sum_{i\in\mathcal{N}} \tilde{v}_i^j w_i^j + \tilde{v}^j z^j = 1, \\ &\sum_{i\in\mathcal{N}} w_i^j \le K z^j, \\ &0 \le w_i^j \le z^j,\ \text{for each } i \in \mathcal{N}. \end{aligned} \tag{73}$$

Therefore, (73) holds if and only if there exist $\boldsymbol{w}^j$ and z^j such that all the constraints in (73) hold and $\sum_{i\in\mathcal{N}} \theta_i^j(p^j v_i^j w_i^j - \alpha_i^j) \ge 0$ holds for all $\boldsymbol{\theta}^j \ge \mathbf{0}$, which is equivalent to $p^j v_i^j w_i^j - \alpha_i^j \ge 0$ for all $i\in\mathcal{N}$. Therefore, (73) is equivalent to that, for any $j\in\mathcal{M}$,

$$\begin{aligned}
&p^j v_i^j w_i^j - \alpha_i^j \ge 0, \text{ for each } i\in\mathcal{N},\\
&\sum_{i\in\mathcal{N}} \tilde{v}_i^j w_i^j + \tilde{v}^j z^j = 1,\\
&\sum_{i\in\mathcal{N}} w_i^j \le K z^j,\\
&0 \le w_i^j \le z^j, \text{ for each } i\in\mathcal{N}.
\end{aligned} \tag{74}$$

That (74) holds for all $j\in\mathcal{M}$ is equivalent to that (70) holds. This completes the proof of Lemma 4. □

We now prove Proposition 5 itself. We first show that if (70) holds, then $\boldsymbol{\alpha}\in\mathcal{A}_{GAM}$. By the first inequality of (70), we have $w_i^j \ge \frac{\alpha_i^j}{p^j v_i^j}$ for all $i\in\mathcal{N}$ and $j\in\mathcal{M}$. Plugging this into the first equality of (70), we have

$$1 - \tilde{v}^j z^j = \sum_{i\in\mathcal{N}} \tilde{v}_i^j w_i^j \ge \sum_{i\in\mathcal{N}} \frac{\tilde{v}_i^j \alpha_i^j}{p^j v_i^j} \text{ for each } j\in\mathcal{M}.$$

Thus, by the first and second inequalities of (70), we have

$$\sum_{i'\in\mathcal{N}} \frac{\tilde{v}_{i'}^j \alpha_{i'}^j}{p^j v_{i'}^j} \le 1 - \tilde{v}^j z^j \le 1 - \tilde{v}^j w_i^j \le 1 - \frac{\tilde{v}^j \alpha_i^j}{p^j v_i^j} \text{ for each } i\in\mathcal{N}, j\in\mathcal{M}.$$

Rearranging the terms, we have

$$p^j \ge \sum_{i'\in\mathcal{N}} \frac{\tilde{v}_{i'}^j \alpha_{i'}^j}{v_{i'}^j} + \frac{\tilde{v}^j \alpha_i^j}{v_i^j} \text{ for each } i\in\mathcal{N}, j\in\mathcal{M}.$$

The first, second, and fourth inequalities and the first equality of (70) imply that

$$\sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{p^j v_i^j} \le \sum_{i\in\mathcal{N}} w_i^j \le K z^j = K\frac{1-\sum_{i\in\mathcal{N}} \tilde{v}_i^j w_i^j}{\tilde{v}^j} \le \frac{K}{\tilde{v}^j}\left(1 - \sum_{i\in\mathcal{N}} \frac{\tilde{v}_i^j \alpha_i^j}{p^j v_i^j}\right) \text{ for each } j\in\mathcal{M}.$$

Rearranging the terms, we have

$$p^j \ge \sum_{i\in\mathcal{N}} \frac{\tilde{v}_i^j \alpha_i^j}{v_i^j} + \frac{\tilde{v}^j}{K} \sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{v_i^j} \text{ for each } j\in\mathcal{M}.$$

Therefore, if (70) holds, we have $\boldsymbol{\alpha}\in\mathcal{A}_{GAM}$.

Next, we show that if $\boldsymbol{\alpha}\in\mathcal{A}_{GAM}$, then (70) holds. Given $\boldsymbol{\alpha}\in\mathcal{A}_{GAM}$, define

$$w_i^j = \frac{\alpha_i^j}{p^j v_i^j} \text{ for each } i\in\mathcal{N} \text{ and } j\in\mathcal{M}, \text{ and } z^j = \frac{1}{\tilde{v}^j}\left(1 - \sum_{i\in\mathcal{N}} \frac{\tilde{v}_i^j \alpha_i^j}{p^j v_i^j}\right), \text{ for each } j\in\mathcal{M}.$$

To show (70), it suffices to show the first, second and fourth inequalities hold because the rest of the constraints hold trivially.

Since $p^j \ge \sum_{i\in\mathcal{N}} \frac{\tilde{v}_i^j \alpha_i^j}{v_i^j} + \frac{\tilde{v}^j}{K} \sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{v_i^j}$ for all $j\in\mathcal{M}$, we have

$$\sum_{i\in\mathcal{N}} w_i^j = \sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{p^j v_i^j} = \frac{1}{p^j}\sum_{i\in\mathcal{N}} \frac{\alpha_i^j}{v_i^j} \le \frac{K}{\tilde{v}^j}\left(1 - \sum_{i\in\mathcal{N}} \frac{\tilde{v}_i^j \alpha_i^j}{p^j v_i^j}\right) = K z^j \text{ for each } j\in\mathcal{M}.$$

Hence, the second inequality of (70) holds. Since $p^j \ge \sum_{i'\in\mathcal{N}} \frac{\tilde{v}_{i'}^j \alpha_{i'}^j}{v_{i'}^j} + \frac{\tilde{v}^j \alpha_i^j}{v_i^j}$ for each $i\in\mathcal{N}, j\in\mathcal{M}$, we have

$$w_i^j = \frac{\alpha_i^j}{p^j v_i^j} \le \frac{1}{\tilde{v}^j}\left(1 - \sum_{i\in\mathcal{N}} \frac{\tilde{v}_i^j \alpha_i^j}{p^j v_i^j}\right) = z^j \text{ for each } i\in\mathcal{N}, j\in\mathcal{M}.$$

Therefore, (70) holds. Hence, the feasible region of $\boldsymbol{\alpha}$ is characterized by (68). This completes the proof. □

G. Resource Allocation and Other Ad-Allocation Problems

For an arbitrary resource-allocation problem, we can reformulate the program $(\mathcal{OTP})$ by setting the cardinality of offer set to 1, and the choice probability ϕ_i^j to 1 when i is offered, otherwise 0, and changing b_i to b_i^j, as follows:

$$\begin{aligned} \max_{\boldsymbol{\alpha}\geq\mathbf{0}} \ & \mathcal{V}_{\mathsf{CT}}(\boldsymbol{\alpha}) \\ \text{s.t.} \ & \sum_{j\in\mathcal{M}} p^j \max_{i\in\mathcal{N}} \theta_i^j - \sum_{i\in\mathcal{N},j\in\mathcal{M}} \alpha_i^j\theta_i^j \geq 0, \ \text{for each } \boldsymbol{\theta}\geq\mathbf{0} \\ & \sum_{j\in\mathcal{M}} b_i^j\alpha_i^j \leq \frac{B_i}{T}, \ \text{for each } i\in\mathcal{N}. \end{aligned} \tag{75}$$

Setting $\theta_i^j = 1$ for some j and all other $\boldsymbol{\theta}$'s equal to zero in the first group of constraints of (75), immediately implies that if $\boldsymbol{\alpha}$ is single-period feasible, we have

$$\sum_{i\in\mathcal{N}} \alpha_i^j \leq p^j, \ \text{for each } j\in\mathcal{M}. \tag{76}$$

Reversely, if (76) holds, then for any $\boldsymbol{\theta}\geq\mathbf{0}$, we have

$$\sum_{j\in\mathcal{M}} p^j \max_{i\in\mathcal{N}} \theta_i^j \geq \sum_{i\in\mathcal{N},j\in\mathcal{M}} \alpha_i^j \max_{i\in\mathcal{N}} \theta_i^j \geq \sum_{i\in\mathcal{N},j\in\mathcal{M}} \alpha_i^j\theta_i^j,$$

i.e. the first group of constraints of (75) holds and it is equivalent (76).

Especially, for the AdWords problem, each vertex $i\in\mathcal{N}$ has budget B_i, and edge (i,j) has a bid b_i^j. When a vertex $j\in\mathcal{M}$ arrivals, we have to match it to a vertex $i\in\mathcal{N}$ who has not yet spent all its budget. After the matching, b_i^j is depleted from B_i. The goal is to maximize the total bid spent. So the formulation is as follows:

$$\begin{aligned} \max_{\boldsymbol{\alpha}\geq\mathbf{0}} \ & \sum_{i\in\mathcal{N},j\in\mathcal{M}} b_i^j\alpha_i^j \\ \text{s.t.} \ & \sum_{i\in\mathcal{N}} \alpha_i^j \leq p^j, \ \text{for each } j\in\mathcal{M}, \\ & \sum_{j\in\mathcal{M}} b_i^j\alpha_i^j \leq \frac{B_i}{T}, \ \text{for each } i\in\mathcal{N}, \end{aligned} \tag{77}$$

and the Display Ads problem, each edge (i,j) has a weight w_i^j, and each vertex $i\in\mathcal{N}$ has capacity c_i. When a vertex $j\in\mathcal{M}$ arrivals, we have to match it to a vertex $i\in\mathcal{N}$ who can be matched at most c_i times. The goal is to maximize the total weight of the matched edges. So the formulation is as follows:

$$\begin{aligned} \max_{\boldsymbol{\alpha}\geq\mathbf{0}} \ & \sum_{i\in\mathcal{N},j\in\mathcal{M}} w_i^j\alpha_i^j \\ \text{s.t.} \ & \sum_{i\in\mathcal{N}} \alpha_i^j \leq p^j, \ \text{for each } j\in\mathcal{M}, \\ & \sum_{j\in\mathcal{M}} \alpha_i^j \leq \frac{c_i}{T}, \ \text{for each } i\in\mathcal{N}. \end{aligned} \tag{78}$$

After solving the above problems and then obtaining the optimal $\boldsymbol{\alpha}$, we can apply the DWO policy to get the allocation. Hence, our DWO policy can be applied to these resource-allocation problems.

H. Implementation Details of the Re-Solving Benchmarks

In this section, we provide the implementation details of the re-solving benchmarks: the Fluid-R, Fluid-I-R, and Fluid-E-R policies. For all policies, we preset a set of re-solving epochs $\mathcal{T} := \{t_u : u = 0, 1, 2, ..., U\}$ in which the algorithm re-solves the Fluid convex program with updated ad budgets and click-through requirements, where $t_0 = 1$ refers to solving ($\mathcal{OP}_{\mathsf{Fluid}}$) at the beginning of the planning horizon. At each re-solving epoch t_u ($1 \le u \le U$), an Fluid-R, Fluid-I-R, or Fluid-E-R policy will re-solve the following convex program with budget and click-through requirement updates (similar to Appendix B, we define $Y_i^j(t) := \sum_{\tau=1}^{t-1} y_i^j(\tau)$ as the cumulative click-throughs until the beginning of time t):

$$\begin{aligned}
\max_{\boldsymbol{z}} \ & \mathcal{V}_{\mathsf{Fluid}}(\boldsymbol{z}|u) := \sum_{i\in\mathcal{N}, j\in\mathcal{M}, S\in\mathfrak{S}^j} r_i^j p^j \phi_i^j(S) z^j(S) + \lambda F(\boldsymbol{\zeta}) \\
\text{s.t.} \ & \sum_{j\in\mathcal{M}, S\in\mathfrak{S}^j} b_i p^j \phi_i^j(S) z^j(S) \le \frac{B_i - b_i \sum_{j\in\mathcal{M}} Y_i^j(t_u)}{T - t_u + 1} \text{ for each } i \in \mathcal{N} \\
& \sum_{j\in\mathcal{C}, S\in\mathfrak{S}^j} p^j \phi_i^j(S) z^j(S) \ge \frac{\eta_i^{\mathcal{C}} - \sum_{j\in\mathcal{C}} Y_i^j(t_u)}{T - t_u + 1} \text{ for each } i \in \mathcal{N} \text{ and } \mathcal{C} \in \mathfrak{K}_i \\
& \sum_{S\in\mathfrak{S}^j} z^j(S) \le 1 \text{ for each } j \in \mathcal{M} \\
& z^j(S) \ge 0 \text{ for each } j \in \mathcal{M}, S \in \mathfrak{S}^j \\
& \boldsymbol{\zeta} \in \mathbb{R}^{nm}, \text{ with } \zeta_i^j = \frac{1}{T} \cdot Y_i^j(t_u) + \frac{T - t_u + 1}{T} \cdot \sum_{S\in\mathfrak{S}^j} p^j z^j(S) \phi_i^j(S).
\end{aligned} \tag{79}$$

It is clear from the formulation that (79) is reduced to ($\mathcal{OP}_{\mathsf{Fluid}}$) with $u = 0$. We denote the solution to (79) in re-solving epoch u as $\boldsymbol{z}^*(u)$. Therefore, the Fluid-R, Fluid-I-R, and Fluid-E-R policies re-solve (79) for U times at the re-solving epochs $\mathcal{T}$, and follows the static policy $\pi_{\mathsf{Fluid}}(\boldsymbol{z}^*(u))$ from time t_u to time $t_{u+1} - 1$ for $u = 1, 2, ..., U$, where we adopt the convention $t_{U+1} = T + 1$.

The only difference between the Fluid-R, Fluid-I-R, and Fluid-E-R policies is the pattern of the re-solving epochs $\mathcal{T}$. More specifically, for the Fluid-R policy (see, also, Jasin and Kumar 2012), $\mathcal{T}$ is evenly spread across all time, i.e., $t_u = \lceil \frac{Tu}{U} \rceil$, where $\lceil \cdot \rceil$ refers to the ceiling function. For the Fluid-I-R policy (see, also, Bumpensanti and Wang 2020), $\mathcal{T}$ is sparser at the beginning of the planning horizon and denser at the end. Specifically, following Bumpensanti and Wang (2020), we set the re-solving epoch $t_u = \lceil T - T^{(5/6)^u} \rceil + 1$ for all $u \in \{1, 2, ..., U\}$, where $U = \lceil \frac{\log(\log(T))}{\log(6/5)} \rceil$. We remark that, in our numerical experiments, we set $T = 1{,}000$ and, therefore, $U = 7$ for the Fluid-I-R policy. To make the benchmark policies more comparable, we set $U = 7$ for the Fluid-R policy as well. Finally, for the Fluid-E-R policy, it re-solves (79) in each period, i.e., $U = 999$ and $t_u = u + 1$ for each u (see, also, Balseiro et al. 2023).

I. Concentration Parameter

In our numerical experiments (Section 6), we vary CP to change the uniformness of proportions $p^1, ..., p^m$ of m customer types, which are generated by a m-dimension Dirichlet distribution. The m-dimension Dirichlet distribution has m concentration parameters $\beta_1, ..., \beta_m$. In our experiments, we set $CP := \beta_0 = \beta_1 = \beta_2 = ... = \beta_m$. Note that, for all j, $\mathbb{E}[p^j] = \beta_j / \sum_{k=1}^{m} \beta_k = \frac{1}{m}$ and $Var(p^j) = \frac{m-1}{m^2(m\beta_0+1)}$, which is decreasing in β_0. For

$j \neq k$, the covariance between p^j and p^k is $-\frac{1}{m^2(m\beta_0+1)}$, which is increasing in β_0. Hence, if β_0 is larger, the sampled customer type distribution will be close to the uniform distribution on $\{1,2,...,m\}$. In contrast, if $CP = \beta_0$ is small, the customer type distribution is more likely to be concentrated on a subset of $\{1,2,...,m\}$. In other words, the higher the CP, the more uniform the generated distribution of customer types.

J. Mean-Reverting Behavior of the DWO Policy

To highlight the mean-reverting property of our proposed DWO algorithm, we examine the intertemporal correlation between of the click-through $y_i^j(t)$ of ad i by type-j customers in period t and the *per-period debt*, defined as $\Delta_i^j(t) := d_i^j(t)/t$ where $d_i^j(t)$ is the debt of ad i for customer segment j at the beginning of time t (as defined in Algorithm 1). Recall that debt measures the gap between the click-through target set by the algorithm and the realized click-throughs. Therefore, if the correlation between $y_i^j(t)$ and $\Delta_i^j(t)$ is larger, it implies that the algorithm "pays back" the "debt" faster and, therefore, the mean-reversion of the click-through process is stronger.

For each of the 5 algorithms studied in our numerical experiments, we regress the click-through on the per-period debt using the following model specification with 30 million randomly drawn samples for each policy studied:

$$y_i^j(t) = a_0 + a_1 \Delta_i^j(t) + \epsilon$$

The regression results are reported in Table 3. If we instead regress the click-through on the total debt $d_i^j(t)$, the results will be similar because the per-period debt is a constant multiplication of the total debt.

Policy	Coefficient	Estimation	Standard Error	t-statistics	p-value
Fluid	a_0	0.0036905	1.1071e-05	333.35	0
	a_1	-0.00042435	0.0015441	-0.27481	0.78346
Fluid-R	a_0	0.0037289	1.1128e-05	335.09	0
	a_1	0.044064	0.0016192	27.213	4.505e-163
Fluid-I-R	a_0	0.0037046	1.1092e-05	334	0
	a_1	-0.0051239	0.0015941	-3.2144	0.0013071
Fluid-E-R	a_0	0.0036979	1.1084e-05	333.63	0
	a_1	0.061641	0.0016548	37.25	1.0702e-303
DWO	a_0	0.0038891	1.1218e-05	346.67	0
	a_1	0.48795	0.0020018	243.76	0

Table 3 The Regression Results of the Intertemporal Correlations Between Click-Throughs and Per-Period Debts

Table 3 shows that our DWO algorithm clearly drives the mean-reverting pattern for the click-through process, captured by the fact that the estimate $\hat{a}_1 = 0.48795$ is positive, large and statistically significant. This is expected given that the DWO policy gives a higher weight for the ad/customer pair with a larger debt at each time t. An important observation from our regression results is that, the estimated coefficient $\hat{a}_1$ of our DWO algorithm (0.48795) is about one order of magnitude larger than that of the Fluid-based benchmarks.

Such an observation delivers an intriguing insight that our debt-based algorithm drives the click-through process toward its mean (i.e., the target set by the first-stage optimization) and, as a consequence, result in a more stable budget depletion process for the ads. Finally, we remark that, because of the budget constraints of the ads, the Fluid-based benchmarks also exhibit certain mean-reverting property weaker than our DWO algorithm.

K. Comparison With the Inventory-Balancing Policies

In this section, we compare our DWO algorithm against another family of benchmarks called the inventory balancing (IB) policies. Specifically, Golrezaei et al. (2014) propose two IB algorithms which implement real-time personalized offer-set optimization with an exponential penalty function (the EIB policy) and a linear penalty function (the LIB policy), respectively, to reweight the value of each ad. Upon the arrival of customer t, the IB policies solve a single-period offer-set optimization problem with a discounted value $r_i\Phi(B_i(t-1)/B_i)$, where $\Phi(\cdot)$ is an increasing discount function and $B_i(t-1)$ is the budget of ad i at the end of time $t-1$. The discount function is $\Phi(x)=(e/(e-1))\cdot(1-e^{-x})$ under the EIB policy and is $\Phi(x)=x$ under the LIB policy. It is hard, if not impossible, to incorporate the non linear fairness metric and the click-through requirements into the IB policies. To account for both the budget constraints and the click-through requirements, one needs to design weight functions handling them jointly. In a case where one ad has little remaining budget, but also falls behind the schedule of its click-through requirements, it is unclear how we should design the weight functions to adjust the weight of this ad. Even worse, the click-through requirements are imposed at the ad by subset of customer types level, which may not be compatible with the ad level budget constraints. Therefore, even without incorporating the fairness metric, it is highly nontrivial to extend the IB algorithm that embeds the click-through requirements. So we remove these modeling features in the comparison between DWO and IB policies. We consider the same numerical setup as Section 6 with the click-through requirements and the fairness metric removed, and the identical per-click value r_i of each ad i across all customer types, sampled from a uniform distribution on the interval $[10,50]$.

We report the results on the comparison between our DWO policy and the EIB and LIB algorithms in Table 4, with the ratio between the standard error of the total advertising revenue for each policy examined to the theoretical upper-bound of advertising revenue included in the parenthesis. The most important takeaway from our experiments is that the DWO policy outperforms the EIB and LIB algorithms when LF is low, especially when $LF<1$. In this case, the budget constraints are not binding, so discounting the ad value when the budget is low is not helpful. On the other hand, when the loading factor is high, the budgets are more likely to be exhausted, so the discount functions of the EIB and LIB algorithms can help smoothly allocate the ad budgets, thus giving rise to higher efficiency performance than the DWO algorithm. To conclude this section, we remark that, because of the difficulty to incorporate the nonlinear fairness metric into the IB policies, this family of algorithms are not amenable to address the algorithmic fairness issue, which can be well handled by our DWO policy.

CP	LF	EIB	LIB	DWO
0.1	1.5	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.25	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.00	99.44% (0.18%)	99.44% (0.25%)	99.20% (0.48%)
	0.75	91.79% (0.21%)	90.66% (0.28%)	99.00% (0.63%)
	0.5	89.19% (0.31%)	86.72% (0.32%)	98.96% (0.49%)
1	1.5	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.25	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.00	99.43% (0.31%)	99.45% (0.24%)	99.40% (0.39%)
	0.75	91.75% (0.17%)	90.60% (0.24%)	98.81% (0.63%)
	0.5	89.23% (0.29%)	86.86% (0.24%)	98.67% (0.64%)
10	1.5	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.25	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.00	99.54% (0.21%)	99.52% (0.27%)	99.37% (0.39%)
	0.75	91.64% (0.17%)	90.53% (0.25%)	99.19% (0.41%)
	0.5	89.20% (0.23%)	86.76% (0.24%)	98.75% (0.51%)
100	1.5	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.25	100.00% (0.00%)	100.00% (0.00%)	100.00% (0.00%)
	1.00	99.59% (0.25%)	99.56% (0.20%)	99.35% (0.40%)
	0.75	91.68% (0.19%)	90.53% (0.21%)	99.29% (0.29%)
	0.5	89.20% (0.18%)	86.71% (0.23%)	98.93% (0.62%)

Table 4 **Numerical Results (Standard Error Relative to the Theoretical Upper Bound in Parentheses)**